The Dissertation Committee for Robert Bradley Shaffer
certifies that this is the approved version of the following dissertation:

Power in Text: Extracting Institutional Relationships from Natural Language

Committee:

Zachary Elkins, Supervisor

Bryan Jones

Stephen Jessee

John Wilkerson

Power in Text: Extracting Institutional Relationships from Natural Language

by

Robert Bradley Shaffer

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2018

Power in Text: Extracting Institutional Relationships from Natural Language

Publication No. _______________

Robert Bradley Shaffer, Ph.D.
The University of Texas at Austin, 2018

Supervisor: Zachary Elkins

How do legislators allocate policy-making authority? At least in the legal context, distribution-of-power arrangements are usually articulated in written documents. Unfortunately, extracting these relationships is difficult, leading scholars to restrict themselves to studies of single policy areas or to a small set of high-visibility laws. In this project, I address this limitation through a neural network-based approach that extracts power relationships from legal language in a scalable, valid fashion. I then apply this approach to study institutional design in enacted US legislation.

Substantively, I demonstrate that policy preferences of executive and legislative actors exert surprisingly little influence on formal institutional design choices. For all but the most politically salient laws, implementation arrangements are structured by the *policy area* and *issue* under consideration rather than elite political preferences. This argument - which would not have been possible to test without the measurement tools I develop - highlights both the importance of the tools I develop and the need for scalable measurement techniques in political science.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Statutory Design in Public and Academic Discourse

How do legislators allocate policy-making authority? As any would-be lawyer knows, statutes, constitutions, and other formal legal texts establish relationships between actors, describing who can do what, when, and to whom. However, the nature and complexity of the institutional structures created by these texts varies enormously. To take a pair of (seemingly) extreme examples from American politics, the Clean Air Act grants the Environmental Protection Agency sole decision-making authority over most policy decisions, while the Patient Protection and Affordable Care Act (ACA) divides decision-making authority between the Departments of Labor, Treasury, Health and Human Services, and Veterans' Affairs, as well as various subordinate agencies and other institutions.

In public discourse, debates surrounding formally-articulated power structures usually reduce to an efficiency/accountability tradeoff. For a concrete example, again consider the ACA. As a core part of Barack Obama's first-term agenda, the ACA represented a sweeping overhaul of the American healthcare system, which expanded availability of public and private insurance plans and empowered regulators to implement an wider set of health insurance regulations. Supporters argued that these changes would protect consumers and would empower expert administrators to increase the overall efficiency of the healthcare system. Detractors, by contrast, worried that the ACA placed undue authority with unelected bureaucrats, who might interfere with individual citizens' healthcare choices. The discussion surrounding so-called "death panels" offers a immediate example of these concerns. Though fantastical, the fear that government regulators might be given the authority to control end-of-life decision-making for Medicare patients resonated with voters, and represented one of the most durable talking points in the entire ACA debate (Oberlander 2012).

As the "death panel" example suggests, both public and academic debates regarding legislative proposals often reduce to arguments regarding the *content* of the text in question. At first glance, basic factual disagreements of this kind might seem surprising. After all, in the United States texts of proposed bills are publicly available as soon as they are formally introduced on the floor of Congress, allowing interested academics or members of the public to peruse their contents. Unfortunately, when dealing with complex and highly-technical documents, access to *text* is not the same as access to *information.* Most legal documents are long, complex, and written with a specialized audience in mind, preventing non-experts from extracting relevant information regarding their topic area or the administrative structures they create. As a result, despite their importance few people actually read legal texts, limiting ordinary citizens' ability to engage with the contents of these documents and leaving members of the media vulnerable to rumors or misinformation campaigns.

1.2 Measurement Limitations as Theoretical Constraints

The academic literature on formal allocation of authority follows similar themes to these public discussions. Like members of the broader public, political scientists and legal scholars generally frame allocation of authority decisions using an efficiency/accountability tradeoff, in which legislators must balance efficiency and responsiveness against oversight over - and protection against - executive malfeasance (e.g. Moe 1990b; McCubbins 1985). Particularly in the American context, scholars have formalized this intuition through versions of the "ally principle" (Epstein and O'Halloran 1999; Huber and Shipan 2002; Bendor and Meirowitz 2004; Franchino 2007; Gailmard and Patty 2012; Farhang and Yaver 2016). When the executive's interests are aligned with those of the legislature (and are likely to remain so in the future), these authors argue, legislators are more willing to pass simple, "framework" legislation, and leave implementation choices to the bureaucracy. By contrast, when legislative and executive interests diverge (or might do so in the future), legislators tend to favor complex institutional structures, which provide greater oversight opportunities and divide decision-making authority among a wider array of implementing bodies.

Beyond these basic political factors, however, we should also expect the design of a given legal document should be affected by characteristics of the *issues* and *policy areas* the document addresses (e.g. Epstein and O'Halloran 1999). For example, we should expect lawmakers to be more willing to design complex, cross-cutting decision-making structures when addressing issues that implicate multiple policy areas or policy areas in which lawmakers possess some preexisting expertise. Work from both the rational choice and bounded rationality literatures suggests that factors such as these should represent some of the most important predictors of formal allocation of authority, which need to be considered alongside preference-related variables in order to fully capture the complexity of the lawmaking process.

Unfortunately, while issue and policy area variables have been discussed extensively in the theoretical literature on formal allocation of authority, ideas like these are largely ignored in empirically-oriented scholarship. Largely, this limitation results from a measurement problem. Even for motivated and well-resourced academics, reading and interpreting legal texts is labor-intensive. As a result, most empirical work on allocation of authority has been restricted to single policy areas or to small sets of "important" legislation. For example, Epstein and O'Halloran (1999) and Farhang and Yaver (2016) only examine legislation on Mayhew (1991)'s list of "important" national-level American legislation, while Huber and Shipan (2002) focus exclusively on bills addressing health policy.

These self-imposed constraints are sensible - and, indeed, necessary, with existing measurement technology - but they are also limiting. As I argue in this dissertation, we should expect formal institutional design choices in most lawmaking scenarios to be shaped by attributes of the *policy area(s)* and *issue(s)* under consideration, rather than the preferences of executive and legislative actors. Preference-oriented explanations, I suggest, help explain institutional design choices contained in "important", high-visibility legislation, but have limited utility outside of this domain.

1.3 A Text-as-Data Approach

In this dissertation, I address these limitations in three parts. First, from a theoretical standpoint, I provide a cognitive microfoundation for many of the existing theories regarding formal institutional design. Broadly, I argue that the process of constructing a legal document is best viewed in terms of a cognitive cost-benefit analysis. When writing a legal text, lawmakers must weigh the policy and political benefits offered by a complex, highly detailed institutional structure against the cognitive costs that designing such structures imposes. Viewed from this perspective, we should expect lawmakers to be most willing to create complex institutional structures when (1) the issue under consideration implicates multiple policy areas and (2) lawmakers possess some preexisting expertise in the policy area in question. For sufficiently important issues, we might also expect preference disagreements between the executive and legislative branches to prompt lawmakers to create more complex institutional structures, but this relationship is likely to be weak in other situations.

Second, I introduce a novel conceptualization and measurement approach designed to help capture the content of formal legal texts. I argue that formal legal documents are best viewed as *relational* documents, which describe which actors can do what, when, and to whom. This network of inter-actor relationships, I suggest, forms what we might call an "implementing network", which describes the institutional content of the document. After describing this conceptual paradigm, I operationalize this scheme using a series of entity and relation extraction techniques drawn from computational linguistics and computer science, and assess its performance on an original dataset of enacted American legislation.

This approach has two major advantages over existing work. From an academic perspective, transforming texts into implementing networks allows researchers to formulate questions regarding the relationship between background political and policy factors and downstream allocation of authority in a coherent theoretical and statistical framework. Network analysis is a mature field, which offers an array of straightforward measures that neatly align with theoretically important concepts such as institutional complexity and fragmentation. Perhaps more importantly, however, this approach also offers substantial descriptive utility. Representing bills as networks of inter-actor relationships offers an intuitive way for non-experts to grasp the basic institutional structures they create. While information of this kind clearly cannot replace expert judgments in a legal or policy setting, this information can help inform public discourse, allowing members of the media and the public to more easily rebut concerns like the ACA "death panel" example I cite at the outset of this section.

Third, from a substantive standpoint I use the techniques I develop to study lawmaking patterns in both American enacted legislation and the US Consolidated Code. In particular, I focus on studying the effects of *policy complexity* and *issue prioritization* on the structure of downstream legislation. Though well-studied in the theoretical and broader policy design literatures (e.g. Epstein and O'Halloran 1999; Bendor and Meirowitz 2004; Jones and Baumgartner 2005; Baumgartner et al. 2009; Baumgartner and Jones 2010; Jones and Baumgartner 2012; Shaffer et al. 2016), due to the data limitations I outline previously most preexisting studies have either ignored the relationship between these ideas and formal allocation of authority, or studied their relationship across a limited set of policy areas or legislative documents. The measurement techniques I develop allow me to address these relationships in a unified modeling context and to study texts addressing low-attention policy areas and issues.

1.4 Limitations of this Study

As I emphasize throughout the preceding discussion, my focus on this dissertation is on *formal* (textually-articulated) distribution of authority. Before proceeding, it is worth pausing to consider the limits imposed by this choice. Clearly, written agreements such as statutes and constitutions do not necessarily represent the actual, on-the-ground distribution of power across a bureaucracy or government, which is determined at least as much by informal practice and circumstances as by formal obligations (see, e.g., Carpenter and Krause 2015). Similarly, these kinds of documents also do not represent either the starting or ending point of the contractual negotiating process for any given institution. Instead, written agreements are best viewed as a function of the incentives, interests, and balance of power among the relevant parties at a particular point in time, as well as the anticipated future state of these features.

Thus, though documents like statutes and constitutions are relevant to the day-to-day functioning of a bureaucracy or a government, they are most directly informative about the political and institutional processes that produced their provisions. As a result, for the remainder of this paper, I focus on questions related to these processes, and present a number of hypotheses and conceptual ideas related to the legislative systems and practices that underlie them.

Importantly, though, this focus on political negotiation should not suggest that formal allocation-of-authority decisions possess little or no practical impact. Clearly, the relationship between formal, legally-articulated power structures and day-to-day practice and outputs is itself an empirical question, which is likely closer or more distant depending on the particular substantive and institutional setting. However, in order to investigate the relationship between text and practice, we first need to understand the text, both in a descriptive sense and in terms of the political and institutional processes that produced that text. As such, though I do not directly address the actual institutional practices and outputs of the governing bodies and agencies I examine, this study is relevant to these institutions, and offers insights regarding the relationship between formal agreements and on-the-ground practice.

1.5 Roadmap

The organization of this dissertation roughly mirrors the three-part organizational scheme I articulate above. In Chapter 2, I begin by outlining the existing work on formal allocation of authority in the political science and legal literatures. I then offer a cognitive microfoundation that unifies and extends many of the ideas presented in existing studies. I focus particularly on highlighting the ways in which this framework should lead us to develop different theoretical expectations for "important" laws compared with their more "everyday" counterparts. Finally, I use this microfoundation to develop expectations regarding the relationships between policy area, preference disagreements, and downstream allocation of authority choices.

In Chapter 3, I build upon this theoretical material by introducing, implementing, and testing a novel measurement scheme designed to capture textually-articulated allocation of authority. I begin by introducing a novel, network-based conceptualization of formal authority. I then implement a neural network-based network extraction approach, which I test on an original dataset consisting of all enacted American legislation passed from 1990-2016. Finally, I present a series of case studies designed to illustrate both the utility and the face validity of the approach I develop.

Starting in Chapter 4, I shift away from theory-building, conceptualization, and measurement, and use the framework I develop in my first two chapters to answer the questions I outline at the outset of this dissertation. In Chapter 4, I focus on the American enacted legislation I use to train and test the measurement approach I describe in Chapter 3. For each bill in the dataset, I extract an implementing network, which I use to generate a measure of *fragmentation* in the bill. I then use a Bayesian hierarchical model to examine fragmentation patterns across the dataset. As predicted, outside of "important", high-salience laws, the relationship between ideological variables and downstream allocation of authority is weak. Rather, most implementing decisions are shaped primarily by characteristics of the *policy area* under consideration.

Finally, in Chapter 5 I use a similar approach to the one I present in Chapter 4 to examine over-time variation in statutory fragmentation in the US Code. Using the tools I develop in Chapters 3 and 4, I generate a dataset consisting of year-on-year changes in fragmentation values for all chapters of the US Code from 1995-2016. I then model these changes using a Bayesian hierarchical model. I find that executive-legislative preference disagreements have a limited impact in this context. After controlling for policy context, fragmentation in the Code increases no faster under unified than divided government, reinforcing the findings I offer in Chapter 4.

Chapter 2

Cognitive Foundations of Formal Power

In the political science and legal literatures, formal (textually-defined) distribution of authority is a key phenomenon of interest (Moe 2012; Carpenter and Krause 2015). Social scientists are often interested in studying the inputs, outputs, and performance of various organizations (e.g. economic performance of a business, or policy choices by a government). Clearly, design of decision-making structures play a critical role in research programs of this sort. Whatever the institution of interest, the allocation of power within that institution plays a critical role in determining the substantive goals it pursues and the efficacy with which it achieves those goals.

Generally speaking, scholars have treated formal allocation of authority as a principal/agent problem, in which legislators are forced to balance efficiency and responsiveness against procedural protections and oversight mechanisms. As Moe (1990a; 1990b) and Moe and Caldwell (1994) note, restrictions on implementer discretion are costly to implement. As a result, policymakers will only restrict executive authority when they have a strong incentive to do so. In most studies, legislative/executive disagreement provides the necessary motivation. When executive and legislative actors disagree on policy preferences, these authors suggest, the legislature should be more willing to constrain executive power by creating restrictive, cross-cutting institutional arrangements.

This preference-oriented theoretical framework is appealing and influential, and has helped to motivate an array of important theoretical and empirical projects. However, it is also limited. As I argue in this chapter, however, work in this domain largely ignores the role of the *policy areas* and *issues* under consideration in a given law. These factors, I suggest, likely exert substantial influence over downstream design of legislation;

The remainder of this chapter proceeds in three parts. First, I provide an overview of the types of language present in legal documents, and emphasize this project's focus on language addressing formal allocation of implementing authority. Second, I outline the existing work on formal allocation of authority in political science, with a focus on the literature's strengths and weaknesses as applied to "everyday" lawmaking choices. Third, I use these ideas to develop a general policy- and issue-centric theory of formal institutional design, which - I argue - is better-able to explain formal institutional design choices contained in both "important" and more commonplace lawmaking activities.

2.1 Conceptualizing Formal Power

When a policymaker writes or modifies a formal policy document - such as a constitution, law, or administrative regulation - that policymaker must choose from a range of possible textual models. At one extreme, she could articulate her objectives in great detail, with measurable policy goals, concrete deadlines and procedural requirements, and complex decision-making procedures requiring approval from a variety of individuals or institutions. At the other, she could write a simple "framework" document, which contains vague policy objectives and delegates most authority to one or a few implementing actors. For a concrete example of these archetypes, contrast the Indian Constitution with its counterpart from the United States. At some 145,000 words, the Indian Constitution is one of the longest in-force national-level founding texts, with detailed information regarding the internal organization of government, individual-level fundamental rights and duties, and national-level ideals and aspirations. The United States Constitution, by contrast, covers only 7,600 words, with few details on any of these topics.

Of course, few legal texts differ as extremely as these. Rather, most documents offer varying levels of specificity across various types of language, with some offering detailed institutional arrangements but vague policy prescriptions and others offering the opposite. Again, for concreteness, consider the following three examples:

Endangered Species Act:[1]

[1] 16 U.S.C. 1533(a)(1-2)

The Secretary [of the Interior] shall [. . .] determine whether any species is an endangered species or a threatened species because of any of the following factors:

1. the present or threatened destruction, modification, or curtailment of its habitat or range;

2. overutilization for commercial, recreational, scientific, or educational purposes.

National Defense Authorization Act of 2004:[2]

The Administrator for Federal Procurement Policy, in consultation with the Secretary of Defense, the Administrator of General Services, and the Director of the Office of Personnel Management, shall develop and implement a plan to ensure that the Federal Government maintains the necessary capability with respect to the acquisition of architectural and engineering services.

Constitution of India:[3]

The executive power of the Union shall be vested in the President and shall be exercised by him either directly or through officers subordinate to him in accordance with this Constitution.

Without prejudice to the generality of the foregoing provision, the supreme command of the Defence Forces of the Union shall be vested in the President and the exercise thereof shall be regulated by law.

These three provisions come from different countries (the United States and India) and different types of legal documents (two ordinary statutes and one national constitution). All, however, offer good examples of the kinds of content present in legal texts. Each excerpt clearly addresses a particular policy issue (environment; arms procurement; executive power and defense) and allocates decision-making authority to one or a collection of actors. The first and second excerpts also at least im-

[2] 41 U.S.C. 1128
[3] Constitution of India (1949), Art. 53(1-2).

plicitly articulate a policy goal (biodiversity protection; maintenance of engineering expertise), while the second provides for oversight over the implementing actor's discretionary authority. Moreover, each provision offers a differing level of specificity on each of these dimensions, with the first offering a detailed set of policy goals but little procedural detail and the second providing the opposite.

Faced with this linguistic variability, scholars have developed a variety of schemes to conceptualize the types of ideas contained within legal texts. McCubbins (1985, 724-729), for example, uses a multilevel scheme, which separates post-hoc oversight and management rules (e.g. regulations regarding oversight hearings or appropriations decisions) from formalized, "structural" language allocating and restricting decision-making powers to particular actors. Within the structural subcategory, he identifies a number of separate sub-types, including institutional rules (i.e. number and identity of actors authorized to act in a particular policy area), procedural requirements, and internal organization rules for relevant organizations. Epstein and O'Halloran (1999, 101) and related authors (Franchino 2004, 2007; Ainsworth and Harward 2009; Oosterwaal et al. 2012) divide much further, identifying as many as 14 separate constraint types in the documents they examine. Finally, Huber and Shipan lump these categories together, measuring policy-specific language rather than specific restriction types (Huber and Shipan 2002; Huber and McCarty 2004).

All of these schemes are useful, and all help categorize the types of implementing language present in legal texts. However, for the purposes of this project, I focus language involving *distribution of authority.* McCubbins (1985, 725-726), in his institutional sub-category, provides a useful summary of these kinds of rules:

> Regulations can be administered through civil or criminal suits in the courts, or through independent commissions or executives agencies, through discretion granted to the president or state and local entities [...] The choice of institutional setting by [a legislature] involves a decision on how much independence [that legislature] wishes to grant the administrators (independence from [legislative] control) and the extent to which other decision-makers [...] restrict or influence the choices of the administrators.

This focus on authority-allocating language - which is closely related to the idea of "fragmentation" explored by Kagan (2009), Biber (2011), and Farhang and Yaver (2016) - narrows researcher attention to a pair of simple questions: within a particular legal text, who is empowered to do what? And, to whom?

Limiting our attention to the institutional language contained in legal texts is necessarily restrictive. However, there are good reasons to believe that the kinds of authority-allocating decisions this emphasis highlights are some of the most important choices that policymakers face during the policy design process. By manipulating the background institutional structure, legislators can force implementing actors to cooperate with other (often hostile) players, altering downstream policy outcomes and restricting implementer discretion. Directive policy language, on the other hand, does not always have such a clear downstream impact. As with all textual provisions, evidentiary requirements, deadlines, and substantive decision-making standards only constrain implementor authority to the extent that they are actually enforced. But, virtually all legal language is compatible with an array of interpretations and implementation styles. Institutional structures can help sharpen these constraints, assigning other actors to monitor, amend, and approve agency policies.

2.2 The Existing Work

2.2.1 Institutional Setting and the Costs of Complexity

As I note at the outset of this chapter, detailed policy language - institutional or otherwise - is not *ex ante* desirable. Complicated power-sharing arrangements that fragment implementing authority across an array of actors dramatically curtails administrative flexibility, reducing implementer responsiveness and promoting policy gridlock. Moreover, detailed authority structures are not costless to create. If a policymaker wants to design a complex policy structure with the intention of accomplishing a policy objective, he or she needs to devote substantial time and attention to the process, which usually involves a substantial quantity of research, expert consultation, and experimentation. In complex policy areas, where the links between policy and outcomes are more uncertain, these problems are particularly severe (Bendor and Meirowitz 2004; Epstein and O'Halloran 1999, 73-75, 84-85). Installing "fire alarms" (McCubbins and Schwartz 1984) designed to delegate oversight and protection functions to courts or interest groups may help somewhat, but "fire alarms" and other *ex post* restrictions frequently ineffective compared with *ex ante* alternatives (Epstein and O'Halloran 1999; McCubbins et al. 1989). Thus, at least in the abstract, policymakers should prefer "framework" legislation that offers one or a few implementing actors substantial discretionary authority.

Here, though, Moe's "politics of structural choice" intervenes. Whenever policymaker and implementor possess different substantive preferences, formal mechanisms of bureaucratic control - institutional or otherwise - become more attractive as a policy strategy, offering protection for policymakers against implementer malfeasance (McCubbins 1985). Moreover, as uncertainty about future electoral outcomes rises, restrictive policies become particularly appealing; if a currently-empowered group is unsure if it will hold public office in the future, that group will likely try to passively "insulate" its programs (de Figueiredo Jr 2002). Insulation-style strategies can take many forms, but an institutionally-oriented approach - in which policymakers create complex, fragmented power structures with individually weaker and more constrained implementers - represents one plausible strategy. Whatever the means, however, insulation strategies are particularly appealing for electorally weak groups and members of public interest organizations, whose advocates are likely to hold elected office only sporadically.

These basic observations have spawned an array of insightful theoretical and empirical work, which offers both policy- and politically-oriented refinements to these initial results. From a policy standpoint, major formal studies (e.g. Bendor and Meirowitz 2004) predict that higher uncertainty and policy issue complexity should be associated with lower institutional complexity and greater delegation of authority. These ideas are sensible, and match predictions we might produce using a cognitivelyinspired frame like that suggested by Baumgartner et al. (2009). From an empirical standpoint, Franchino (2004, 2007) and Huber and Shipan (2002) find that better-informed legislatures tend to delegate less, lending support to this idea. Unfortunately, as I describe in the following section, little empirical work in this area cuts across multiple policy areas, leaving few opportunities to examine the relationship between policy complexity and downstream allocation of authority in detail. Epstein and O'Halloran (1999), in an exception to this trend, do examine bills covering an array of policy areas, and find support for these hypotheses; however, their study is limited to Mayhew (1991)'s list of "significant" legislation, limiting the scope of their results.

The relationship between political variables and downstream allocation of authority has been developed in a more extensive fashion. Various authors have, for example, found that legislative capacity, the availability of informal restrictions on delegation, and expectations of future electoral success are all significantly related to allocation of authority (Epstein and O'Halloran 1999; Franchino 2007; Ainsworth and Harward 2009; Huber and Shipan 2002). For the most part, projects in this area have operationalized preference disagreements using simple party control of the executive and legislative branches. However, at least in the US context, there is some evidence that the majority party's seat share also matters (Epstein and O'Halloran 1999; Farhang and Yaver 2016).

One particular focus in this body of scholarship is the relationship between executive/legislative preference disagreements and downstream allocation of authority. As

I describe previously, the so-called "ally principle" - that principals (usually legislators) should be more comfortable offering discretionary authority to their ideological allies than their enemies - is a key motivating idea throughout the statutory design literature. Moe and Caldwell (1994)'s comparison of lawmaking practices under presidential and parliamentary government provides perhaps the sharpest articulation of this idea. In classical parliamentary system like the United Kingdom, the head of government is elected (and removed) by the ruling party (or coalition) in the legislature. As a result, the executive and legislative branches in these countries should pursue similar policy goals, leading legislatures in these countries to pass simple, "framework"-style legislation. By contrast, they suggest, presidentialism-style independent election of the head of government should create more preference divergence between the executive and the legislature, leading the legislature to pass more detailed and restrictive laws.

Theoretical predictions regarding the impact of preference disagreements become more complicated when multiple principals or implementing agents are present. Broadly, when multiple "bosses" possess mutual veto power over a particular policymaking choice, the relationship between preference disagreements and downstream distribution of authority depends on the agenda-setting powers of each principal (Bendor and Meirowitz 2004) and the policy expertise of the agency tasked with implementing the policy (Volden 2002; Gailmard 2002, 2009). From a public administration standpoint, Gailmard and Patty (2007) and Gailmard and Patty (2012) incorporate bureaucratic independence and expertise into this basic modeling framework, and reach mixed conclusions regarding the relationship between preference disagreements and policy design choices and outcomes when a choice of implementing agents is available. And, in a related stream of work, Bertelli and Feldmann (2006) and Prendergast (2007) find that incorporating constituent preferences and electoral pressure into the basic allocation-of-authority model leads policymakers to select "biased" implementing actors, who can credibly commit to implement a more favorable policy outcome than that preferred by their political opponents.

Unfortunately, empirical verification of this latter set of predictions also relatively limited. Epstein and O'Halloran (1999), in their study of Mayhew (1991)'s list of "significant" American legislation, find that delegation to independent agencies and commissions increases under divided government, while delegation to politically-controlled agencies increased under unified government. Huber and Shipan (2002) report a similarly contingent relationship, with a negative relationship between divided government and delegation in professionalized state legislatures and a positive one in non-professionalized settings. Like many empirical studies in this area, Huber and Shipan (2002)'s efforts are limited to a single policy area - namely, state-level Medicaid reform - limiting the generalizability of their results. However, combined with related studies, their findings are at least suggestive of a more complicated, contingent relationship between preference disagreement and downstream allocation of authority than that suggested by early work in this area.

2.2.2 The Cognitive Foundations of Lawmaking

Based on this body of work, we now possess strong knowledge regarding the relationship between formal allocation of authority and a variety of political and policy factors. However, despite these advances, significant gaps remain. Most prominently, as I emphasize throughout the previous section, little empirical work has compared patterns in allocation of authority across multiple policy areas. The few exceptions to this rule (e.g. Epstein and O'Halloran 1999; Farhang and Yaver 2016) are limited in an equally important fashion. In particular, these studies only include prominent or "significant" legislative proposals, without examining their lower-profile counterparts.

These limitations, I suggest, are substantial, and impose serious constraints on our ability to understand lawmaking activity as a general phenomenon. In institutions like the US Congress, most enacted laws are not the kinds of high-profile and transformative legislative actions that make lists of historically important legislation. For example, take Mayhew (1991)'s list of "significant" legislation, which is perhaps the most prominent example of its kind. In his description, Mayhew characterizes his list as a compendium of the most innovative and consequential laws passed by Congress.[4] As a result, Mayhew's list is short - covering some 218 laws enacted between 1946 and 2008[5] - and mostly populated by the kinds of memorable, high-visibility laws already familiar to scholars of legislative behavior.

Though by definition important and worthy of study, laws of this kind are clearly not

[4]See, e.g., Mayhew's notes accompanying the 2015 update to his list.

[5]As updated by Farhang and Yaver (2016).

representative of the bulk of lawmaking activity. Most obviously, because of their brevity lists of this kind must inevitably make decisions about edge cases, such as the Clean Water Act[6] (included on Mayhew's list) or the Endangered Species Act[7] (excluded). However, borderline cases like these are rare compared with the more "everyday" laws that make up the bulk of the lawmaking output in most contexts. For a concrete example, take the recent history of the United States Congress. From 1989-2008 Congress enacted some 4,186 unique pieces of legislation. Of these, approximately one-quarter (1,018) were relatively trivial "commemorative" bills, which established monuments or symbols designed to commemorate noteworthy people or places in US history. Some 63 further laws from this period made Mayhew's list of historically noteworthy legislation. The remaining three-quarters, however, were somewhere in between, neither entirely trivial nor significant enough to merit particular historical mention.

Perusing these laws in the middle category reveals a predictably mixed set of legislation. Unsurprisingly, some "everyday" bills are essentially inconsequential, and consist only of technical amendments to - or extensions of - existing elements of US law.[8] Some, however, are more impactful. For a concrete example, consider the Enhanced Partnership with Pakistan Act (2009).[9] This bill authorized some $1.5 billion in yearly defense-related aid to the government of Pakistan from 2010-2014. Though

[6]Pub. L. 92-500.

[7]Pub. L. 93-205.

[8]E.g. Pub. L. 108-306, "To provide an additional temporary extension of programs under the Small Business Act and the Small Business Investment Act of 1958 through September 30, 2004, and for other purposes".

[9]Public Law No. 111-73.

not significant enough to make a list of historically significant legislation, this bill nevertheless involved substantial financial outlays and represented an important shift in US-Pakistan relations. It also contained some noteworthy administrative innovations, the most notable of which was an unusual oversight system that gave the State Department primary responsibility over allocation of funds (Epstein and Kronstadt 2013).

Laws like the Enhanced Partnership with Pakistan Act form the bulk of the lawmaking outputs in most contexts. As a result, any theory that seeks to explain the lawmaking process as a *general* phenomenon needs to be able to explain these "everyday" bills as well as their higher-salience counterparts. Importantly, existing work in American politics and public policy gives us good reason to suspect that drafting processes underlying "significant" laws should look substantially different from their "everyday" counterparts. For example, consider Baumgartner et al. (2009)'s work on policy attention:

> As a given social indicator becomes more troubling over time, the [bounded rationality] model predicts no response whatsoever during the early periods; the issue is "under the radar" and government may not even track its severity in any systematic manner. After the severity of the issue has passed some threshold, on the other hand, there may be a rush to make up for past inattention to the issue by dramatically increasing policy outputs directed to it. The issue may be systematically tracked and a specialized agency or bureau may even be created to focus on it. (Baumgartner et al. 2009, 608)

Applied to the formal institutional design context, this framework produces a natural set of predictions. Because lawmakers possess limited staffing and cognitive resources, we should not expect them to develop precise policy preferences for most policy issues. Consequently, we should also not expect lawmakers to propose institutional arrangements designed to express their (vague or nonexistent) preferences in all scenarios. Instead, we should only expect lawmakers to express their policy preferences through formal institutional design choices when addressing the kinds of important, high-attention policy issues described by Baumgartner et al. (2009). Importantly, as I discuss earlier in this chapter, the relationship between lawmaker preferences and downstream institutional design is frequently a complex and contingent one. Much of the existing work in the literature on formal institutional design focuses on the ways in which institutional and political context condition lawmaker incentives when legislative and executive actors disagree on policy preferences. However, whatever the direction of this relationship for a given law or policy problem, we should expect the link between lawmaker preferences and downstream institutional design choices to be small or negligible for "everyday" policy problems.

If lawmaker policy preferences do not offer a general explanation for variation in institutional design choices, what factors are more influential? I argue that characteristics of the *policy area(s)* under consideration offer the answer. Broadly, the policy area(s) implicated by a particular law should affect the institutional structures it contains through two primary mechanisms, which we might describe as lawmaker *experience* with the policy issue(s) under consideration and the inherent *complexity* of those issues.

Beginning with the former set of relationships, we should generally expect lawmakers to be more willing to create complex institutional structures when they possess *more experience* with the issues under consideration. This prediction follows naturally from the broader cognitive ideas that underlie the bounded rationality literature. In educational testing, a central idea in many studies is the notion of "cognitive load", or the "demands on working memory [...] intrinsic to the material being learned" (Paas et al. 2003). In this context, tasks which involve difficult or unfamiliar subject matter (e.g. abstract literary criticism; complex logical reasoning) impose a higher cognitive load on the subject than tasks which involve easier or more familiar topics (simple mathematical computations; memorization) (see, e.g. Sweller 2010).

Translated to the political domain, this basic cognitive pattern suggests that we ought to expect politicians to generate more complex institutional structures in policy domains in which they have more *experience*, since the "cognitive load" involved in researching and designing complex systems in these areas will be lower. Since policy experience is an individual-specific phenomenon, we should expect the cognitive load imposed by a particular policy problem to vary similarly by individuals. However, we should expect most lawmakers to possess more experience with familiar, frequently-occurring issues (e.g. agricultural subsidies; banking regulation) than when addressing less-familiar concerns (e.g. cybersecurity; space and science policy). As a result, lawmakers should generally be more willing to create complex institutional structures in the former set of areas than the latter.

Besides these individual-specific effects, we should also expect lawmakers to create more complex institutional structures when the issue under consideration involves cross-cutting policy concerns. Abstractly, some policy problems are inherently "broader" than others. For example, take the Dodd-Frank Act[10]; though the law was primarily intended as a financial regulation bill, it also affected housing and agriculture policy through various reforms to the mortgage and agricultural futures markets, respectively. By contrast, bills like the Clean Air Act[11] or Clean Water Act[12] focus almost exclusively on environmental policy. Intuitively, this variation in issue complexity should be reflected in institutional design choices. If a law addresses a policy problem that cuts across many policy areas, we should expect bureaucratic agencies and other actors that represent each of those policy areas to be involved in the execution of that law. By contrast, for laws addressing more straightforward policy problems, we should expect to observe a smaller set of implementing actors.

Put together, these ideas offer a powerful theoretical framework from which to understand formal institutional design choices. Generally speaking, implementing arrangements contained in legal texts should be shaped by the characteristics of the *policy issues* addressed by those texts. If lawmakers are more familiar with the policy areas under consideration or if the issues raised by a particular text cut across multiple policy areas, we should expect the text to contain a more complex implementing structure. For particularly "important" or "high-salience" laws, we should also expect lawmakers to spend the time and cognitive energy necessary to craft

[10] Pub. L. 111-203.
[11] Pub. L. 88-206.
[12] Pub. L. 92-500.

institutional arrangements that will produce their preferred policy outcomes. However, because lawmakers possess limited time and attention, the relationship between political preferences and downstream institutional design should be attenuated for more "everyday" lawmaking tasks.

2.2.3 Measurement and Selection Bias in Existing Work

The ideas I describe in the preceding section have not gone entirely unnoticed in existing work. For example, Bendor and Meirowitz (2004) predict that politicians should produce simpler institutional arrangements in simpler policy areas, a proposition for which Epstein and O'Halloran (1999) and Koop (2011) provide some empirical support. Unfortunately, measurement constraints have prevented existing work from addressing the two constraints (policy area and policy salience) highlighted in the previous section. The reason for this limitation is simple: *measuring text-based distribution of authority is difficult.*

To demonstrate the scope of the problem, contrast two of the most prominent allocation-of-authority measures in the literature: specifically, Huber and Shipan (2002)'s word-count measure and Epstein and O'Halloran (1999)'s discretion index. Huber and Shipan, for their part, make a simple argument. When faced with two statutes that address the same issue, they suggest, the longer one typically places greater limits on the actions of other actors than the shorter one (Huber and Shipan 2002, 45). Using this observation, they operationalize restrictions on distribution of authority using document word counts.

This measure is trivial to calculate, but theoretically problematic. Most obviously, as Huber and Shipan (2002) themselves note, length only represents a plausible measure of of discretionary restrictions when comparing legal texts written within a single institutional context that address the same policy issue. Outside of these kinds of comparisons, length may be affected by institution-, issue-, or period-specific drafting conventions, introducing substantial measurement error. However, even within a suitably restricted comparison set, word counts remain problematic as a measure of policy detail. In general, a legal text can be lengthy for two reasons: (1) it may address a single idea in a highly detailed and complex fashion, or (2) it may address many ideas in simplistic fashion (see, e.g. Maltzman and Shipan 2008). This problem is less dramatic when comparing legal texts within a single policy area, but even within the same policy domain, most legal documents are highly multidimensional, and cover a variety of aspects of a given policy problem.

By contrast, Epstein and O'Halloran (1999)'s measure suffers from the opposite set of problems. In their study, Epstein and O'Halloran use Congressional Quarterly's year-end legislative summaries to count the number of major provisions (scope) contained in their statutes of interest. They then count the number of constraints placed on the exercise of those powers, and use those two values to calculation each statute's "discretion index" d_i, operationalized as:

$$d_i = r_i(1 - f_i)$$

Where f_i is the number of constraint types (out of 14 identified by the authors) present in statute i, and r_i is the proportion of major provisions in statute i that offer discretionary authority to the executive (Epstein and O'Halloran 1999, 100-108).

This approach is intuitive and has proven useful in a variety of institutional contexts, both in its original form (see, e.g. Franchino 2004, 2007) and coupled with small modifications or additions (e.g. Koop 2011; Farhang and Yaver 2016). Unfortunately, this approach is also highly labor-intensive. Farhang and Yaver (2016), for example, read and coded some 24,000 pages of legislative text in order to produce data on some 218 laws passed from 1947 to 2008. Epstein and O'Halloran (1999) rely on legislative summaries rather than reading legislative texts themselves; however, this reliance on third-party information produces the selection problem I highlight in the previous section, since only "significant" legislative texts receive the kind of third-party attention necessary for this scheme.

This limitation has not gone entirely unnoticed in recent methodological work. For example, using Franchino (2007)'s data as a training set, Anastasopoulos and Bertelli (2017) train a random forest classifier to predict a version of Epstein and O'Halloran (1999)'s "delegation index" using European Union legal texts. In their paper, the authors report reasonably high out-of-sample performance, with approximately 75% accuracy when predicting the presence of provision-level restrictions on discretionary authority. However, their approach still leaves substantial room for improvement, and their existing results offer little guidance regarding generalization to other institutional contexts.

2.3 Hypotheses

In this project, I fill the gaps I identify previously by examining two basic groups of relationships: specifically, *policy area effects*, and *attention effects*. These ideas can be roughly summarized as follows:

1. *(Policy)* How do allocation of authority decisions vary based on policy area? Are legislators more likely to restrict implementor authority in "familiar" issue areas (e.g. economics or civil rights) than in areas outside their core competences (e.g. energy or science)? And, are the relationships between distribution of authority and previously-identified factors like legislative-executive preference disagreements conditioned by policy area?

2. *(Attention)* How do allocation of authority decisions vary based on issue salience? Are legislators more willing to delegate broad policy authority to implementing actors in low-salience policy areas? Or, are institutional design choices on a given issue largely independent of the level of public attention that issue receives? And, as with the previous set of questions, does issue salience condition the relationship between central ideas like legislative/executive mistrust and downstream allocation of authority?

Intuitively, we should expect both of these factors to substantially affect legislative decision-making. For a stylized example, contrast a pair of (hypothetical) bills modifying the responsibilities of the National Science Foundation (NSF) and Consumer Financial Protection Bureau (CFPB), respectively. The NSF bill addresses a low-salience policy problem located in a policy area with which most lawmakers have little experience. As a result, lawmakers working on this bill would likely be inclined to write a simple, "framework"-style law, rather than investing the energy required to write a more complex piece of legislation. By contrast, the CFPB bill addresses a high-salience issue (consumer financial regulation) in a more familiar policy area, incentively lawmakers to draft a more complex and more fragmented piece of legislation.

Importantly, I do not argue that policy and attention factors are the only variables that should affect downstream design of legislation. Classic variables identified by the existing literature - including divided/unified government or other kinds of executive-legislative preference disagreements - also affect legislative decision-making in important ways. However, insights drawn from the bounded rationality tradition suggest that policy and attention factors are likely to be some of the most important predictors of legislative behavior, both directly and by conditioning the effects of other factors. Unfortunately, as I document in the previous section, measurement limitations have prevented scholars interested in formal allocation of authority from incorporating policy and attention factors into their models of legislative behavior, particularly in empirically-oriented scholarship. Re-focusing the empirical literature on these factors therefore represents one of the core theoretical contributions of this dissertation.

To fix intuition, I provide a formalized statement of hypotheses tested in the empirical chapters of this dissertation. By construction, the statements of these questions

I give are somewhat vague. In particular, I omit most of the details regarding operationalization of key variables until I introduce my two empirical studies in Chapters 4 and 5. I make this choice for two reasons. First, from a theoretical standpoint, the ideas I introduce in this chapter are intended to apply across institutional and legal setting. Since the operationalization choices I make in this dissertation are necessarily shaped by my context of interest (specifically, national-level American legislation), introducing information regarding implementation choices at this stage would detract from the cross-contextual nature of the theoretical points I raise in this section. Second (and perhaps more practically), the operationalizations of most concepts I use in this dissertation rely on a novel network-based conceptualization and measurement approach, which I introduce in more detail in Chapter 3.

H1: All else equal, legislators should create simpler authority structures in complex policy areas outside their core competences than when addressing more familiar and straightforward policy problems.

This idea follows directly from the bounded rationality literature. If a policymaker possesses less experience working in a particular policy area, creating legislation designed to address problems within that area will require greater cognitive effort. For example, lawmakers will need to devote more time and staff resources soliciting expert advice and reading policy recommendations in order to craft complex decision-making structures in unfamiliar areas than in familiar ones. Since policy experience varies by individual, the "cognitive load" imposed by a particular legislative proposal will vary similarly, but we might reasonably expect the average cognitive load for a member of Congress to be lower in familiar, high-profile policy areas like finance or criminal policy than in more technical ones like space and science regulation. As mentioned in the previous section, Epstein and O'Halloran (1999) provide some support for this idea, noting that legislators tend to delegate more on high-complexity policies (e.g. space policy). But, existing empirical work has not tested this idea beyond the small set of historically significant legislation these authors examine.

H2: All else equal, legislators should create simpler authority structures when addressing low-attention policy problems than when examining their higher-salience counterparts.

The individual-level foundations for this hypothesis are very similar to those I offer for **H1**. On average, low-salience policy issues attract little legislative attention, since legislators are usually disinclined to exert the cognitive effort required to create complex decision-making structures without constituent pressure. As a result, when addressing low-attention policy issues legislators should be more likely to delegate decision-making authority to one or a few implementing actors, instead of creating a complex and interconnected decision-making structure.

H3: If the executive and legislative actors possess similar preferences, legislative outputs should contain simple decision-making structures. By contrast, if legislative and executive actors have different preferences, decision-making structures may be either simple or complex.

As described in previous sections, the relationship between legislative/executive preference conflicts and downstream allocation of authority is complex. However, as (Volden 2002, 112) notes, "both the executive and the legislature do indeed have an interest in increasing bureaucratic discretion when their preferences align".[13] This idea - which is further developed by Bendor and Meirowitz (2004) - suggests that set of desirable decision-making structures is relatively small when executive and legislative preferences align, since both principal and agent favor simple authority structures. By contrast, when the legislative and executive disagree, outcomes are less predictable, and decision-making structures may be either simple or complex depending on the characteristics of the implementing agent (and in particular, whether the implementer is independent or subject to political control by the executive).

H4: The relationship between executive/legislative preference conflict and downstream allocation of authority should be attenuated when the policy proposal in question addresses complex or low-attention policy areas.

This idea - which represents an *interactive* relationship - also follows naturally from the theoretical foundations that underly **H1** and **H2**. Legislative and executive actors must both expend resources (cognitive and otherwise) in order to create complex authority structures. On high-salience policy issues in familiar substantive areas, both sets of actors will be highly incentivized to pay these costs. However, when addressing low-salience, complex policy problems, the costs incurred by creating a

[13]Other studies have extended this idea to include expectations regarding *future* preference alignment. For simplicity, I omit considerations of this kind both in this section and in the empirical chapters of this dissertation. However, incorporating expectations regarding future electoral performance represents a direction for future work.

complex decision-making structure will likely overwhelm any policy and electoral payoffs that a higher-quality policy proposal may offer. As a result, in these situations the magnitude of the relationship hypothesized in **H3** will likely be small or nonexistent.

2.4 Conclusion

In this chapter, I make two major arguments: one theoretical, and one measurement-related. On the theoretical side, though formalized allocation-of-authority decisions have received substantial scholarly attention across a variety of institutional contexts, substantial gaps between empirical and theoretical work in this area remain. In particular, scholars have not fully explored the relationships between *issue salience*, *policy area*, and allocation of authority in empirical work, despite their centrality in the bounded rationality literature elsewhere in political science and public policy (see, e.g. Baumgartner et al. 2009).

These gaps, I suggest, are largely due to measurement limitations in existing work. Existing empirical scholarship addressing formal allocation of authority generally relies on one of two approaches: (1) labor-intensive hand-coding schemes, or (2) word-count methods like those suggested by Huber and Shipan (2002). Hand-coding methods, for their part, require a substantial time and labor investment, making comparisons across many different policy areas and across low- and high-salience legislation impractical. Word-count methods are more efficient, but do not generalize easily to cross-contextual comparisons comparisons. Even within a particular legal system, drafting conventions within a particular policy area may lead laws addressing some issues to be more verbose than their counterparts in other areas, regardless of their authority-granting content. In order to fill the theoretical gaps highlighted above, then, we need a more comprehensive conceptualization and measurement strategy.

Chapter 3

From Text to Networks: A Relational Conception of Formal Power

As I suggest throughout Chapter 2, in order to answer many of the key questions developed in the literature on formal institutional design, we need a more scalable and generalizable measurement procedure. In this chapter, I propose a conceptualization and measurement approach designed to address these limitations, which I apply to extract allocation-of-authority information from enacted American legislation. This focus on the national-level American context offers two advantages. First, from a substantive point of view, most existing work on formal allocation of authority focuses on the American context, allowing me to build more directly on existing predictions and theoretical work. Second, from a practical standpoint, national-level United States legislative texts are easily available and consistently organized. This organizational consistency, in particular, substantially simplifies some of the measurement tasks I outline in this chapter.

The remainder of this chapter proceeds in three parts. First, I provide a detailed articulation of the relational conceptualization of formal power I use in this dissertation. As I discuss in Chapter 2, for the purposes of this dissertation I argue that legislative texts are best viewed as *relational* documents, which describe which actors can do what, to whom. This relational conception of formal power lends itself naturally to a network-based approach, in which legal documents are represented as networks of decision-making and authority-allocating relationships. Second, I implement a neural network-based approach designed to extract implementing networks from legal texts, which I validate using a dataset consisting of all American legislation enacted since the early 1990s. Third, I offer several short case studies designed to fix intuition and emphasize the descriptive utility of this approach.

3.1 From Text to Networks

As I discuss at the outset of Chapter 2, my focus in this dissertation is on measuring and studying *institutional* language contained in legal texts. Laws, constitutions, and administrative regulations contain many features that are potentially relevant for allocation-of-authority choices, such as prescriptive policy language or complex procedural requirements. However, as I argue, there are good reasons to believe that the institutional content contained in legal texts has the greatest impact on downstream policy outcomes. Directive language - for example, setting acceptable pollution rates or evidentiary requirements in criminal cases directly in the the text of a law or regulation - is only constraining to the extent that it is actually enforced, which requires cooperation from implementing actors. By contrast, shifting the decision-making structure embedded within a particular legal document offers a more direct way to alter policy outcomes and implementing styles, making this kind of "institutional" language a more theoretically appealing object of study.

Besides its substantive advantages, narrowing our attention to the institutional language contained in legal texts also offers a natural conceptualization and measurement approach. When reduced to their institutional content, legal texts can be viewed as *relational* documents, which describe the set of actors involved in the execution of a particular policy program and the relationships between them. Or, phrased more succinctly, focusing on their institutional content recasts legal texts as *networks* of power relationships between actors situated in a particular policy space.

Treating texts as a network of relationships is common in other substantive contexts. In political science, a notable example of this kind of approach is GDELT (Leetaru and Schrodt 2013), which mines news accounts for subject/action/object triples corresponding to international events. Though not immediately presented in network form, relational representations of this kind can be naturally converted into a network representation. For a more direct example, consider Franzosi et al. (2012)'s "quantitative narrative analysis" approach. In their paper, Franzosi et al. use newspaper accounts of slave lynchings to measure patterns of *agency* within a corpus of Georgia lynching accounts from 1322 newspaper stories written from 1875-1930. Like Leetaru and Schrodt (2013), Franzosi et al. identify and code subject/action/object triples in their corpus, and record usage rates for various actors (African American victims; white citizens; law enforcement) and action types (violence; coercion; search; apprehension). They then collect these triples into a network representation, which they use to visualize and analyze the action types present in the accounts they examine. This approach reveals some surprising findings; for example, at least according to newspaper accounts, law enforcement officials were targeted for coercion by mobs nearly as frequently as African Americans, though the overwhelming majority of of violent actions was directed at lynching victims (Franzosi et al. 2012, 12).

This approach is similarly helpful in the legal setting. From a theoretical perspective, if we view legal texts as networks of power relationships, many substantively important questions reduce to simple network-theoretic hypotheses. For example, work by Epstein and O'Halloran (1999), Volden (2002), Bendor and Meirowitz (2004), and others on relative allocation of authority to independent commissions and politically-controlled administrative agencies can be viewed as a question regarding the importance of each actor type to their respective implementation networks, operationalized through node centrality or influence metrics. Similarly, policy fragmentation (Kagan 2009; Farhang and Yaver 2016) can be operationalized using quantities like the number of nodes or average degree of the network. Better still, because these quantities can be calculated for any network, these kinds of quantities are straightforwardly comparable across any set of networks derived from some formal legal corpus of interest.

From a more practical standpoint, a relational conceptualization of formal power also helps to pinpoint the measurement challenges involved in a study of this kind. At their most fundamental level, networks are constructed from *nodes* (the actors involved in the network) and *edges* (the ties between them). Treating formal legislative texts as implementing networks implies a natural measurement problem: in particular, how best should the relevant nodes and edges be extracted from formal legislative documents? Many approaches to this problem are possible, but for the purposes of this dissertation I turn to a series of tools drawn from the natural language processing (NLP) literature. NLP is a catch-all term describing a set of computational methods that attempt to analyze the linguistic attributes of a given text. NLP methods are thus extremely wide-ranging, covering topics like part-of-speech tagging, grammatical parsing, lexical co-occurrence, latent content analysis (e.g. topic modeling), and much else besides. Coupled with a deep, politically- and legally-informed understanding of the documents in question, these tools can provide a powerful approach to the analysis of legal texts.

3.2 Nodes and Edges: Extracting Textual Implementing Networks from American Law

3.2.1 Constructing the Dataset

As mentioned at the outset of this chapter, my primary application of interest in this dissertation is enacted national-level United States legislation. To study this corpus, I therefore constructed an original dataset consisting of all enacted national-level legislation texts and metadata available through congress.gov, the official U.S.

Congress legislative database.[1] I describe the metadata matching and structural parsing components of this data collection effort more extensively in my substantive chapters; however, at a high level, I use Selenium, Beautiful Soup, and base Python HTML libraries to scrape legislative texts from the congress.gov website. I then stripped leading and trailing administrative content from each document (e.g. date of passage; legislative history; transcription notes), and segmented each document into its organizational components.

For segmenting purposes, I rely on the regular expression parser I implemented in constitute_tools, a set of utilities I designed to assist with Constitute's data collection efforts. This parser separates each document according to a given set of organizational headers (e.g. titles; sections), while maintaining the internal hierarchy of each document (see Appendix A for details). By default, constitute_tools further segments documents into paragraphs (separated by newline characters). However, since most newline characters in congress.gov's plain-text document transcripts result from typographic formatting rather than meaningful substantive choices, for the purposes of this analysis I simply recombine paragraphs for each parsed text.

An example output drawn from the USA PATRIOT Act is given in Table 3.1. This sample reveals some of the complexities of the task I describe in this chapter. In this section, the law establishes an immigration monitoring program targeted at foreign students, and assigns joint responsibility to the Attorney General and Secretary of State for its implementation. However, besides this institutional content, the

[1] The tools used to construct this database are available via Github as Legislative_Data.

Table 3.1: Sample parsed document

Title	Text
SEC 416	FOREIGN STUDENT MONITORING PROGRAM.
(a)	Full «NOTE: 8 USC 1372 note.» Implementation and Expansion of Foreign Student Visa Monitoring Program Required.– The Attorney General, in consultation with the Secretary of State, shall fully implement and expand the program established by section 641(a) of the Illegal Immigration Reform and Immigrant Responsibility Act of 1996 (8 U.S.C. 1372(a)).
(b)	Integration «NOTE: 8 USC 1372 note.» With Port of Entry Information.–For each alien with respect to whom information is collected under section 641 of the Illegal Immigration Reform and Immigrant Responsibility Act of 1996 (8 U.S.C. 1372), the Attorney General, in consultation with the Secretary of State, shall include information on the date of entry and port of entry.

USA PATRIOT Act §416(a-b). For original text see the corresponding congress.gov page.

text also contains an array of transcription notes and references as well as other proper nouns and related language which need to be distinguished from the text's institutional references.

3.2.2 Entity Extraction

Entity extraction is a classic NLP problem, which has been attacked using a variety of heuristic and machine-learning approaches. In political science, perhaps the most common approach is a dictionary-based system (see, e.g. Leetaru and Schrodt 2013), in which the entities of interest are pre-identified using a dictionary generated by expert researchers. This approach generally produces few false negatives, but misses a large number of items of interest, since generating a comprehensive named entity dictionary is impractical in most situations. By contrast, applying standard off-the-shelf entity extraction tools (such as the Stanford CoreNLP parser, as described by Manning et al. (2014)) will likely capture most entities of interest, but will also include irrelevant entries (e.g. names of people or places).

For the purposes of this dissertation, I split the difference between these two approaches by training a specialized entity extraction algorithm. In particular, for estimation purposes I rely on a particular implementation of a long short-term memory recurrent neural network (LSTM). Broadly, neural networks are a class of machine learning approaches which seek to predict some outcome of interest based on a series of "hidden layers", which iteratively manipulate some observed set of predictor variables in order to produce predictions. Recurrent neural networks - of which LSTMs are a variant - build on this framework by allowing the predicted output for a given data point to be influenced by the predictor variables and corresponding predicted outputs for nearby data points, creating a recursive, context-sensitive prediction structure ideal for analyzing sequentially-organized information. Since most linguistic data possesses a natural ordering and context, LSTMs are a natural choice in this domain, and have been used for tasks such as language modeling (Sundermeyer et al. 2012) and part of speech tagging (Huang et al. 2015; Plank et al. 2016).

From an implementation standpoint, I rely on the neural network architecture proposed by Lample et al. (2016) and Ma and Hovy (2016).[2] Given a textual excerpt

[2]As implemented in Tensorflow and Python by Guillaume Genthial. The implementation pro-

(e.g. a sentence or paragraph), this implementation predicts token-specific tags based on three sources of information:

1. pre-trained embedding vectors for each word (here, drawn from GloVe, trained on the Google News corpus and described by Pennington et al. (2014));
2. concatenated character embeddings for each character contained in the word (trained during model estimation); and
3. embedding vectors and predicted tags for left- and right-adjacent terms.

This approach surfaces relevant information for both simple and complex tagging rules. For example, incorporating character-specific information allows the model to easily learn that most named entities begin with a capital letter, while incorporating predicted tags for adjacent words allows the model to correctly tag multi-word entities. Word embeddings, by contrast, incorporate more subtle information regarding word usage and semantic patterns, which can be used to identify words which are commonly contained in institution names of interest (e.g. "Secretary" or "Agency").

To generate training data for this model, I used a two-step procedure. First, I built a custom dictionary of institution names that are likely to be present in American legislative texts. To build this dictionary, I first scrape all names contained in usa.gov, the Federal Register, or one of five Wikipedia sources: specifically, the lists of federal agencies, defunct federal agencies, House committees/subcommittees, Senate committees/subcommittees, and joint committees. I then removed common prefixes

duced by Genthial is slightly different from the one outlined in the two papers I cite in-text; for details, see the accompanying blog post.

and suffixes from these items (e.g. "United States"; "USA"), and stripped names of states and national governments (e.g. "Texas"; "California"; "Federated States of Micronesia") from the list. As an additional quality control measure, an undergraduate research assistant read a random sample of legislative texts, and supplemented this list with a series of additional missing items. The final dictionary produced by this process contained some 1360 items, representing most prominent institutions contained in the executive and legislative branches.

Second, using the dataset described in the previous section, I constructed a set of example sentences that contained entities identified in my entity dictionary (see 3.2 for an example). Using the dataset described in the previous section, I extracted all sentences contained in the body of each legislative text.[3] Next, for each sentence I conducted a simple string search for each named entity contained in my entity dictionary. If a particular entity was present in a particular sentence, I marked the first token of the entity with a "B-MISC" tag (denoting the beginning of the named entity), and any subsequent tokens with an "I-MISC" tag (denoting words inside the named entity). Finally, I marked all tokens not identified using one of these labels with an "O" tag.[4] Since the computational complexity of the LSTM model I use scales with the length of the longest input sentence, to ease computation I then discarded all sentences longer than 75 words. This process left me with a training

[3]Using the pretrained Punkt sentence tokenizer, available via NLTK.

[4]Since some named entities are substrings of others - for example, compare "Secretary of Defense" with "Assistant Secretary of Defense" - before searching each sentence I ordered the named entity dictionary from longest tag to shortest, to ensure that the longest present named entity would be tagged first.

Table 3.2: Sample training example

Token	Tag
Funds	O
herein	O
appropriated	O
to	O
the	O
Department	B-MISC
of	I-MISC
Defense	I-MISC
for	O
construction	O
shall	O
be	O
available	O
for	O
hire	O
of	O
passenger	O
motor	O
vehicles	O
.	O

Sample output, formatted according to the CoNLL2003 format. Military Construction Act 1992 §102. For original text see the corresponding congress.gov page.

set consisting of some 29,080 sentences.

Using this set of training examples, I trained the LSTM model described previously, and assessed its performance.[5] Machine learning models are generally assessed using a combination of several performance criteria. Two common such metrics are

[5]Details regarding hyperparameter specification are given in Appendix A.

precision and *recall*, defined as:

$$P = \frac{TP}{TP + FP}$$
$$R = \frac{TP}{TP + FN}$$

Where P and R denote precision and recall, $T(P|N)$ denote the count of true positive/negative examples correctly classified and $F(P|N)$ denote the counts of true positive/negative incorrectly classified. In the entity extraction setting, *precision* can therefore be interpreted as positive predictive value, or the proportion of extracted entities that actually represent agency or institution names of interest. Similarly, *recall* can be interpreted as the true positive detection rate, or the proportion of total institution names of interest actually extracted by an algorithm.

Most entity extraction methods perform better on one of these standards than the other. For example, returning to the two methods I describe at the outset of this section, dictionary methods offer a high *precision* (few false positives) but a low *recall* (many false negatives). By contrast, off-the-shelf named entity extraction methods offer the reverse, with a high *recall* (few false negatives) but a low *precision* (many false positives). A common evaluation metric for binary classification problems that assesses both of these criteria simultaneously is the F_1 score:

$$
\begin{aligned}
F_1 &= 2\frac{1}{\frac{1}{P}+\frac{1}{R}} \\
&= 2\frac{P \times R}{P+R}
\end{aligned}
$$

Where P and R denote precision and recall as defined at the outset of this section. The F_1 score therefore represents the harmonic mean of precision and recall, and offers a reasonable way to balance between these two criteria.

How well does the LSTM approach perform by this standard? Since the purpose of using a machine learning approach for named entity recognition is to capture named entities not already known to the researcher, the most relevant (and stringent) performance test would be one in which we assess the model's ability to recover unseen entities not available during testing. In order to assess the model's performance in this scenario, I therefore conducted a cross-validated predictive accuracy study. In particular, I first randomly split my entity dictionary into five equally-sized groups. Beginning with the first group, I identified all sentences exclusively containing entities from the group in question, and used these sentences to form a held-out test set. I then trained a model using sentences containing entities from the remaining four groups, predicted values for the held-out test set, and used these predictions to calculate predictive accuracy and F1 score values. Finally, I repeated this process for each group in the dataset, and averaged the performance statistics to produce my final results.

Assessed in this fashion, the LSTM model I employ achieved a cross-validated F1 score of 0.758, with an overall accuracy of 0.965. Since negative examples (non-named entities) are much more common than positive examples, (named entities), this difference between F_1 score and overall accuracy is not surprising; as with any imbalanced class problem, the model can simply guess the common class (here, the "O" tag, which denotes the absence of a named entity) and raise its overall classification accuracy. By contrast, the F_1 score focuses our attention on both the common and rare classes, and offers a more reasonable evaluation metric.

How should we interpret these results? Using a near-identical approach to the one I employ, Lample et al. (2016) report an F1 score of .904 on the standard CoNLL2003 named entity test dataset while Ma and Hovy (2016) report an F1 score of .912 and an overall accuracy of .976, both of which are noticeably better than the results I report. However, the performance test I use in this paper is also noticeably more stringent than those used in other studies. In most studies of this kind, researchers assess performance by splitting *sentences* into training and test sets, rather than splitting *entities* into training and test sets. As a result, any given entity can (and usually does) occur in both the training and the test sets, creating a substantially simpler measurement ask. As a result, though the performance statistics I report leave some room for improvement, since they approach the values offered in other work I suggest that they represent a strong starting point from which to work.

3.2.3 Relation Extraction

Compared with entity extraction, relation extraction is a substantially more complex problem. Identifying a particular word or phrase as a named entity involves analyzing some data about that word or phrase (and perhaps its local context), and reaching a classification decision. By contrast, analyzing the *relationship* between two entities involves analyzing the entities, their local context, and any words or phrases which might encode information regarding their relationship. This additional information increases the scope of the problem substantially, since it requires the model to consider larger blocks of text (e.g. words separating two named entities) and, for most applications, a larger typology of relationship types.

Given a predefined set of training examples for a particular relation types, researchers have had some success in generating approaches designed to extract these relationships, though performance results are usually substantially lower than in the named entity extraction case (e.g. Roth and tau Yih 2004; Surdeanu et al. 2011). Unfortunately, for the purposes of this project, generating training data corresponding to all conceivable relation types is not practical. For a project of this kind, we would need expert human readers to locate and code pairwise relationships between all actors based on a preexisting typology of relation types, such as oversight or joint decision-making relationships. Though exciting as a direction for future research, when combined with the existing data collection and coding efforts already present in this project, generating textual networks based on a large typology of relation types is too substantial a coding task for this project.

Fortunately, as I describe in Chapter 2, many variables of interest in the literature on formal allocation of authority do not rely on a detailed typology of relation types. For example, take (Farhang and Yaver 2016)'s work on fragmentation of authority. In their study, they define fragmentation as a tripartite concept, which counts the *number* of *distinct* actors empowered to execute a *particular statutory provision.* Though this definition might be enriched by a more nuanced definition of actor types or relationships between them, a more abstract notion of "fragmentation" still offers the authors rich theoretical ground with which to work. As (Farhang and Yaver 2016) argue, decisions regarding the optimal level of fragmentation implicate a version of the efficiency/accountability tradeoff I raise at the start of this dissertation; in their view, fragmented implementation arrangements are inefficient and promote incoherent/contradictory policy outcomes, but also "insulate" agencies from capture and create institutional fire alarms and other accountability-enhancing mechanisms.

For both practical and theoretical purposes, I therefore focus on an abstract tie type, which is similar to that identified by Farhang and Yaver (2016). In particular, I define a tie between two actors as an instance in which two actors are *assigned to implement the same policy program.* Luckily, drafting guidelines for American legislation make these kinds of relationships relatively easy to identify. As noted in the drafting guide for the US Consolidated Code, the "basic unit" of every section of Code and legislation is the *section.*[6] Laws and Consolidated Code fragments are often further subdivided into ordered list elements of various types, but *sections*

[6]http://uscode.house.gov/detailed_guide.xhtml

are intended to be stand-alone units that are roughly comparable in substantive scope. As a result, if we observe that two actors are co-mentioned in a section of a law, we can reasonably conclude that those two actors share authority over the policy area under consideration in that section. Without a sharper definition of the relationships under consideration, we cannot draw strong conclusions about the nature of the connections between these actors, but we can draw general conclusions about the basic implementation structure envisioned by the law in question.

The relation extraction procedure I employ in this paper, then, proceeds in two steps. First, I segment each text according to its internal organization (e.g. titles, sections, etc), and remove preliminary material, section titles, and section headers from the text.[7] Second, I recombine each text into sections, extract entities from each section, and draw an edge between any set of entities that co-occur in a given section. These two components combine to form the extracted "implementing network" for each law, which forms the basis for the other analyses I present.

3.3 Applications

Before concluding this section, I examine two brief case studies to highlight the descriptive utility of this approach. The two cases I provide here - specifically, the Enhanced Partnership with Pakistan Act of 2009 and the America Recovery and Reinvestment Act of 2009 - are intended to demonstrate both the promise and the

[7] Using the parser contained in https://github.com/rbshaffer/constitute_tools.

challenges involved in the network extraction methodology I present in this chapter.

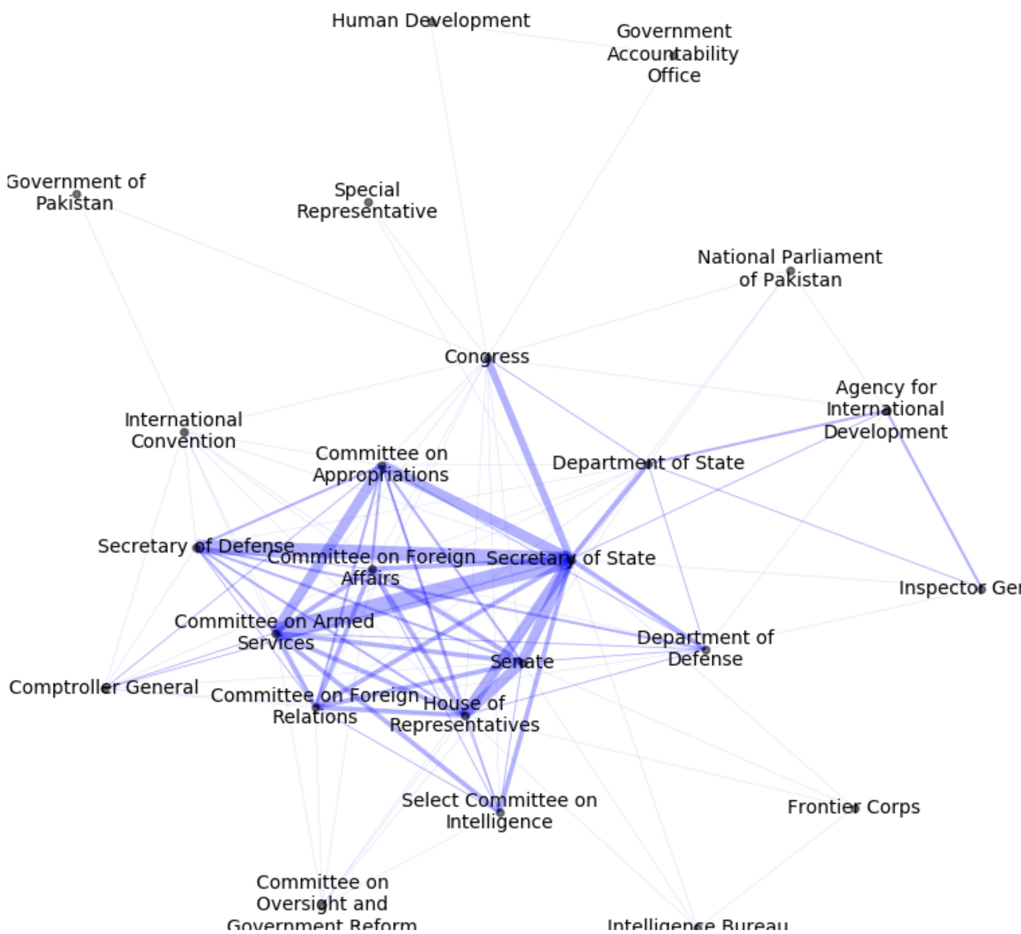
Beginning with the simpler case, the Enhanced Partnership with Pakistan Act of 2009[8] is a relatively straightforward foreign aid bill intended to provide military and developmental assistance to the Pakistan. The law authorized the President to provide $1.5 billion in non-military aid from 2010-2014, and provided additional military aid conditional on a certification process implemented by the Secretary of State.[9] From an administrative perspective, the law established two funds - specifically, a Counterinsurgency Fund and a Counterinsurgency Capability Fund - administered by the Defense Department and State Department, respectively (Epstein and Kronstadt 2013). Unusually for a defense-oriented bill, the law gave the State Department substantial authority over defense-related aid allocations Epstein and Kronstadt (2013).

All of these features are clearly visible in the implementation network for this law, which I show in Figure 3.1. The law contains a central cluster consisting of the Secretary of State, the Secretary of Defense, and several Congressional actors, including the House and Senate floors, the Appropriations Committees (both chambers), the Armed Services Committees, and the House Committee on Foreign Affairs, and the Senate Committee on Foreign Relations. Quantitative assessments of node importance reinforce this visual message; for example, as assessed by eigenvector centrality, the Secretary of State is the most central actor in this network (eigenvector centrality

[8]Public Law No. 111-73. For original text see corresponding congress.gov page.

[9]Enhanced Partnership with Pakistan Act of 2009, §203.

Figure 3.1: Authority network extracted from the Enhanced Partnership with Pakistan Act of 2009



Line density is approximately proportional to the number of ties between each node. Node placement is random, but is loosely related to node centrality.

of 0.46), followed by the the Committee on Armed Services (0.37), and the House and Senate floors (0.33). These statistics roughly track qualitative summaries of the bill's organization; in both cases, the State Department emerges as the most prominent institutional actor in this network, with other US government actors occupying a less central role.

From a face validity standpoint, the visualization shown in Figure 3.1 shows both the promise and the challenges of this approach. As expected, the neural network-based named entity extraction model I employ identifies virtually all of the primary actors of interest from this law, including several (e.g. "National Parliament of Pakistan"; "Pakistan Counterinsurgency Fund") not included in the named entity training set used to fit the model. However, the model also identifies some false positives (e.g. "Pakistan Assistance Strategy Report"; "Human Development") and improperly merges (House/Senate Appropriations Committees) some names. Overall, though, these cases are relatively rare, and occupy an outsized position in Figure 3.1 because of the visualization's tendency to isolate rare nodes at the plot's edge.

In contrast with the Enhanced Partnership with Pakistan Act, the American Recovery and Reinvestment Act (ARRA)[10] is both broader in scope and substantially more complex in its institutional organization. Briefly, the ARRA is a stimulus bill designed to bolster American economic performance following the 2007-2008 Financial Crisis. The bill's provisions - which can be roughly divided between infrastructure investments, tax cuts, and direct fiscal assistance to state and local governments.

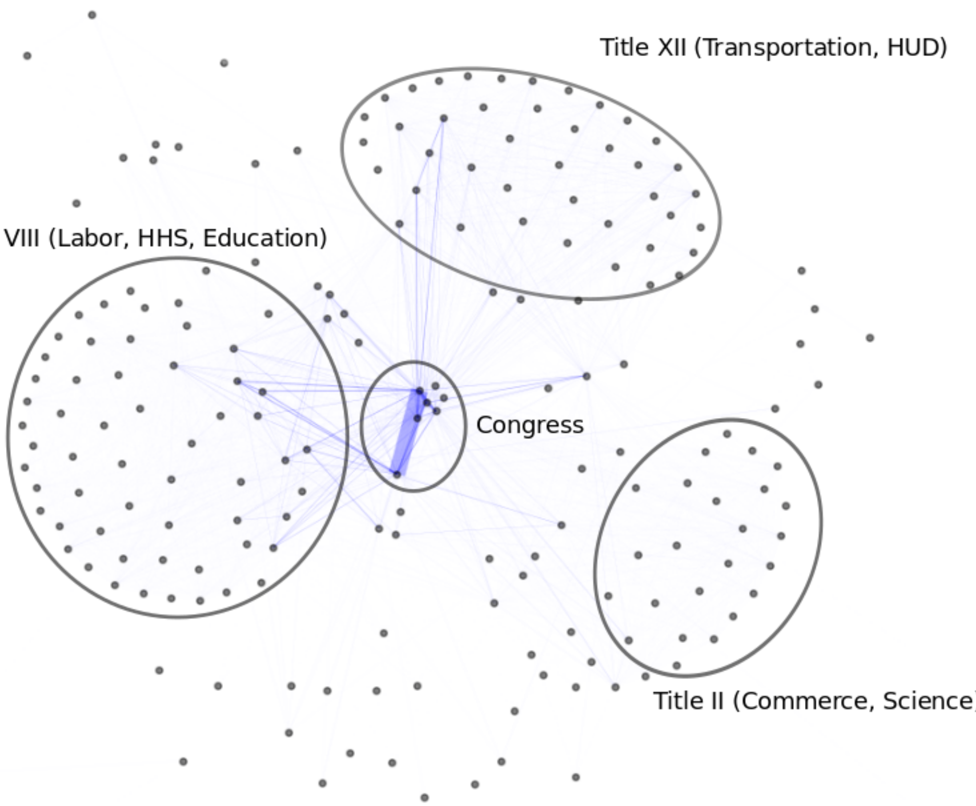
[10]Public Law No. 111-5. For original text see corresponding congress.gov page.

Within the fiscal assistance component of the bill, the ARRA allocated the largest portion of its funds via Medicaid, food stamps, unemployment compensation, and other programs administered through the Department of Health and Human Services (Auerbach et al. 2010). Healthcare-related spending also formed a large part of the ARRA's investment spending, though infrastructure programs formed the bulk of the bill's spending in this area.

Compared with the Enhanced Partnership with Pakistan Act, the full implementation network for the ARRA is substantially more complex. Since the ARRA covers so many policy areas and programs, this complexity is reassuring, but it makes interpretation substantially more difficult. Upon inspection, however, some clear structure emerges from the network. As before, we can begin by examining statistics generated using this network to learn something about its overall composition. For example, as measured by eigenvector centrality, the most central nodes of this network are the Senate (0.44), House (0.39), and Inspector General (0.30). Among executive branch nodes, the Secretary of Labor (0.15), Department of Defense (0.14), Secretary of Homeland Security (0.13), and Secretary of Health and Human Services (0.13) appear as the most central nodes.

The node placement algorithm I use to generate Figure 3.2 provides another way to interpret the ARRA's structure. Roughly speaking, the graph visualization algorithm I employ in Figure 3.2 places nodes with a greater number of connections between them in close proximity to one another. As a result, a simple visual inspection of the plot produced using this method reveals that the bill contains a central

Figure 3.2: Full implementation network for the American Recovery and Reinvestment Act of 2009



Line density is approximately proportional to the number of ties between each node. Node placement is random, but is loosely related to node centrality.

cluster corresponding to Congress (e.g. the House/Senate and various Congressional committees). Institutions covered by larger titles of the bill - such as Title VII (covering the Departments of Labor, Health and Human Services, and Education) and Title XII (Transportation and Housing and Urban Development) - form roughly coherent visual clusters, which are placed at the outer edges of the plot. This structure is reassuring; since the ARRA is roughly organized by executive department, we should expect the bill's implementing network to display a roughly coherent set of clusters, which correspond to the major arms of the bill.

Comparing the overall organization of the ARRA network with its counterpart from the Enhanced Partnership with Pakistan Act is similarly revealing. For example, one idea we might be interested in studying is the *overlap* between the administrative structures created by each of these laws. Alternatively, this idea could viewed as the degree of "siloing" present in each law, with a highly "siloed" law delegating large discretionary authority to individual actors and an interconnected one offering a more fragmented institutional structure. One way to operationalize this idea is to examine the *average clustering coefficient* of the network. This statistic (which is defined in the binary setting as the average number of closed triangles passing through each node) measures the extent to which nodes in the network are densely connected. Compared across networks, the clustering coefficient will be lowest when many isolated nodes, branches, or groups are present in the network.

Examining the clustering coefficients of these networks gives an average clustering coefficient of 0.195 for the Enhanced Partnership with Pakistan Act, compared with

0.016 for the ARRA. These values possess substantial face validity. From a qualitative standpoint, the ARRA clearly possesses the wider scope of the two, authorizing more policy programs and addressing a broader set of policy areas. However, as a consequence of this breadth, the implementation network shown in Figure 3.2 is much sparser than its counterpart from Figure 3.1. Phrased differently, the programs contained within the ARRA are largely independent, with most actors restricted to their particular areas of expertise. By contrast, the Enhanced Partnership with Pakistan Act separates implementing authority for its programs across several actors, creating a more interconnected decision-making structure.

3.4 Conclusion

Overall, then, the network extraction method I propose in this chapter proceeds as follows. First, I segment each text of interest according to its internal hierarchy. For American legislation, drafting guides dictate that the relevant unit of analysis is the *section*, but other texts might use other approaches. Second, using a custom-trained recurrent neural network, I extract all institution names from each unit. These names form the *nodes* in the network I extract. Third, I draw an edge between any two institution names that co-occur within a particular section, forming a network of relationships between institutions of interest.

As I demonstrate, this approach performs reasonably well across the dataset of American legislative texts I examine in this section. For the named entity extraction component of this method, out-of-sample performance results are roughly similar to those reported by other users of the LSTM model I employ. The implementation networks produced in my two case studies also show strong face validity, with network statistics generating using these networks matching qualitative descriptions of their content. Though performance in all of these areas is certainly not perfect, these results are encouraging, and offer some reassurance regarding the quality of the outputs produced by the method.

Chapter 4

The Design of US Legislation

Up to this point, the material in this dissertation has been focused on conceptualization, theory construction, and measurement. In this chapter, I return to the key empirical question I introduce in Chapter 2: specifically, how do legislators allocate policymaking authority in formal legislative texts? In particular, I focus on studying *fragmentation of authority.* As I describe at the outset of this dissertation, when designing a formal legislative texts, lawmakers can choose between one of two basic models. On the one hand, they can create a simple "framework" law, which provides few detailed policy goals and allocates authority to one or a few implementing actors. Alternatively, they can create a complex and detailed piece of legislation, which divides implementing authority between many actors in a complex fashion. Existing literature suggests that the choice between these options should be a function of preference disagreements between the executive and the legislature; however, as I argue in Chapter 2, this relationship is likely to be moderated by issue salience and policy area. The method I develop in this dissertation offers a natural way to measure design of legislation in a scalable fashion, allowing researchers to test these kinds of relationships without resorting to labor-intensive hand-coding schemes.

The remainder of this section proceeds in three parts. First, using the method described in Chapter 3, I operationalize fragmentation from a network-theoretic perspective. Second, I introduce a dataset consisting of all enacted American legislation passed since the early 1990s, which I use to test the hypotheses introduced previously. Third, I describe a Bayesian hierarchical model designed to test these hypotheses, which I then implement and interpret. As predicted, I find high-salience bills passed under divided government are substantially more fragmented than their lower-salience, unified-government counterparts. However, the effects of both of these variables vary substantially by policy area, emphasizing the importance of the measurement and modeling techniques I develop.

4.1 Overview

In the study on formal institutional design, *fragmentation* of decision-making authority is a key concept of interest. Broadly speaking, institutional design decisions can be viewed as invoking tradeoffs between efficiency and accountability. Institutional structures that require more precise outcomes, impose more onerous procedural requirements, or create more complex decision-making structures are generally less efficient, but offer greater opportunities for lawmakers to hold implementing actors accountable to the preferences of the broader public. Fragmenting implementing authority offers a straightforward example of this kind of tradeoff. By dividing policymaking authority between several actors, lawmakers reduce implementing efficiency and limit implementor discretion, but offer more opportunities for outside groups to monitor and intervene into the policymaking process. "Siloing" each individual implementor, by contrast, insulates implementing actors from other agencies and from scrutiny by outside groups, offering an opposite tradeoff. Because institutional rules are self-implementing we should expect implementors to be especially responsive to variation in fragmentation of authority, making these kinds of rules an especially fruitful area of study.

I describe my theoretical expectations regarding fragmentation patterns in more detail in Chapter 2. Briefly, however, I suggest that institutional design choices can be best explained by weighing the cognitive costs of constraint against their political and policy payoffs. Creating complex decision-making mechanisms is costly for legislators, who must invest time, fiscal resources, and mental energy in order to create them. Legislators should therefore be most willing to pay these costs when they believe fragmented authority structures will be easier to create, when the electoral payoff for creating those structures will be high, or when they are particularly suspicious of executive malfeasance.

I summarize these intuitions using four hypotheses. First, all else equal, legislators should create simpler authority structures in complex or unfamiliar policy areas. Institutional design choices in more complex policy areas demand greater cognitive resources, disincentivizing legislators from creating complicated and fragmented policy structures in these domains. Second, legislators should create simpler authority structures when addressing low-salience issues. In these domains, claiming credit for political achievements is more difficult, reducing lawmaker incentive to invest the time and effort necessary to create complex authority structures. Third, legislators should create simpler authority structures when the legislative and executive branches agree on policy preferences. And, finally, the relationship between executive/legislative preference conflict and downstream allocation of authority should be attenuated when the policy proposal in question addresses low-attention issues.

4.2 Assembling the Bills Dataset

To study empirical fragmentation patterns in enacted American legislation, I return to the training dataset I used in Chapter 3 to fit the named entity extraction neural network I employ in this dissertation. This dataset consists of all enacted legislation for which both text and metadata are available through congress.gov - the official U.S. Congress legislative database - and the Congressional Bills Project. I then filtered commemorative bills from the dataset[1], and constructed an implementing network for the remaining texts using the procedure outlined in Chapter 3. Briefly, I first segmented each text according to its internal organization (e.g. titles, sections, etc), and removed section titles, and section headers from the text.[2] Next, I removed the first section from each document. In contemporary American legislation, the first section of each document always contains a set of preliminary material, such as an

[1]As identified by the Congressional Bills Project.

[2]Using the parser contained in https://github.com/rbshaffer/constitute_tools

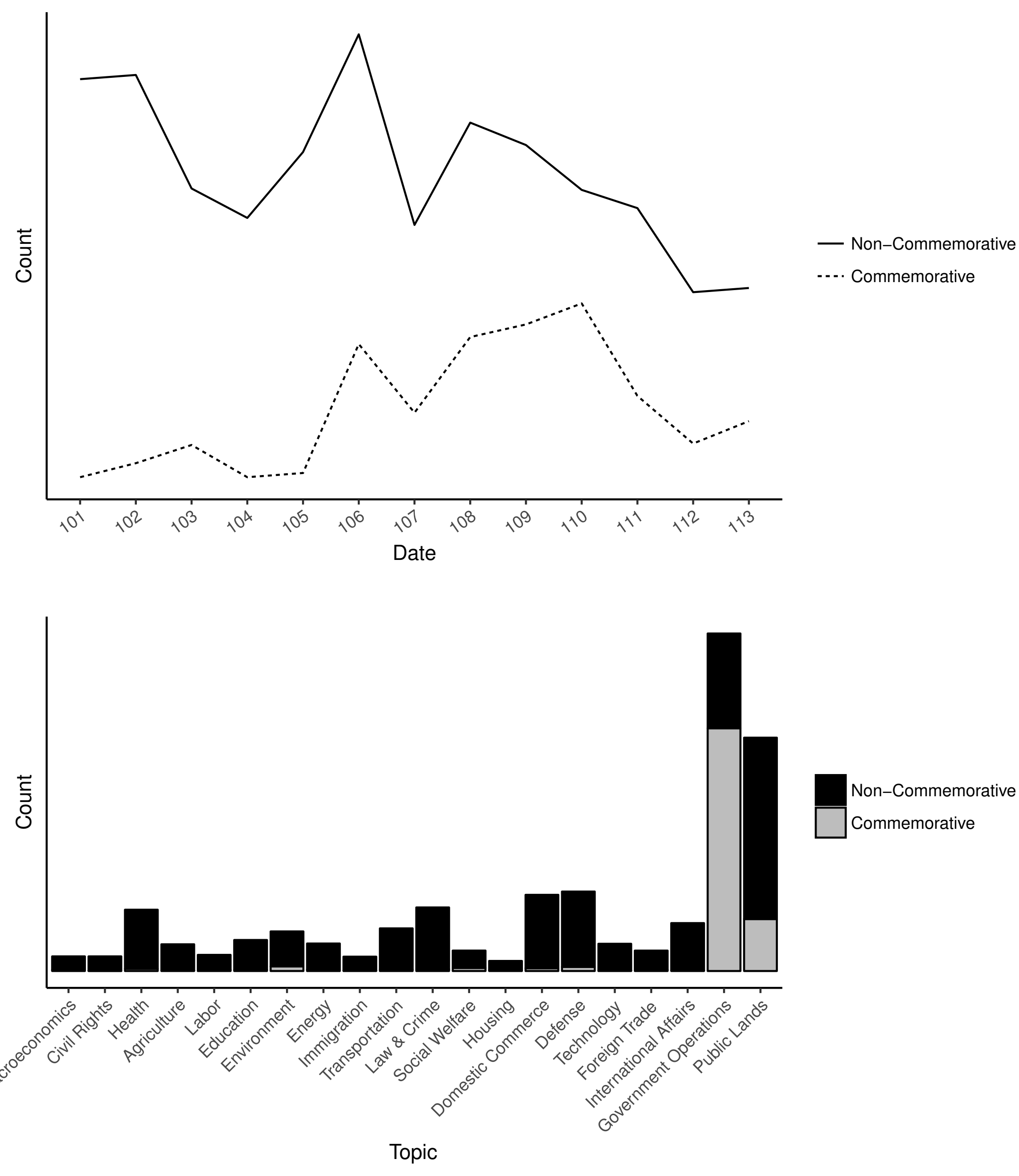
official "short title", a table of contents, or similar information, which is not relevant for the analytical task I undertake in this chapter. I then extracted entities using the pre-trained LSTM model I describe in Chapter 3, and discarded all named entities that were mentioned only once in their respective documents.[3] Finally, I recombined each text into its constituent sections, and drew an edge between any set of entities that co-occur in a given section.

Descriptive information for this dataset is given in Figure 4.1. Temporally, this dataset approximately covers the period from 1990-2014. At the earlier part of this period, congress.gov's coverage is not complete. As a result, conclusions drawn from this period should be interpreted with caution. Across the dataset, some 3467 observations (out of 4800 total) are both covered by the Congressional Bills Project's metadata and non-commemorative.[4] As shown in Figure 4.1, the most common topics in the dataset (both with and without commemorative bills included) were public lands and government operations; however, all major topics (as identified by the Policy Agendas Project) are represented.

[3]This heuristic is drawn from the natural language processing literature (see, e.g., Grimmer and Stewart 2013), and is useful in cases where typographic errors or other types of false positives are likely to be common.

[4]Additionally, 42 observations from other parts of the dataset were missing due to metadata errors in the Congressional Bills Project's data.

Figure 4.1: Descriptive information for the enacted American legislation dataset



4.3 Variables

4.3.1 The Dependent Variable: Network Fragmentation

Despite its theoretical importance, relatively few authors have operationalized fragmentation in a systematic fashion. McCubbins (1985) and Kagan (2009), for example, offer broad overviews of bureaucratic organization and types of constraints on implementor discretion, but little in the way of empirical operationalization of fragmentation or related ideas. Farhang and Yaver (2016) offer perhaps the most systematic description of the concept, defining "fragmentation" as:

> Division of implementation authority over a larger number of distinct actors, over a larger number of different agencies, and giving multiple actors the authority to perform the same function with respect to the same statutory provisions.

Fragmentation, from this perspective, therefore refers to the *number of actors involved in the execution of any particular statutory function.* Farhang and Yaver (2016) operationalize this idea a fairly direct fashion. For each statute they examine, they read the statutory text and code (1) the number of actors empowered to execute a core regulatory function in the law, (2) the number of federal agencies empowered to execute a core regulatory function in the law, and (3) the number of instances in which two or more actors or agencies were given simultaneous authority over a particular policy area. They argue that these ideas all tap aspects of their underlying concept of interest (which they further support using visual and statistical evidence), and therefore use a simple average of these three components as the final dependent variable in their analyses.

Translated to the network context, this three-part conceptualization scheme offers a natural starting place from which to operationalize fragmentation of implementing authority. As Farhang and Yaver (2016) note, both a simple count of the number of actors/federal agencies tasked with implementing key policy programs and the number of instances in which actors are offered overlapping policy authority within a particular statute offer reasonable proxies for the extent to which authority within that statute is fragmented. As a result, either of these measures offer a reasonable proxy for the level of fragmentation present in a particular law.

Fortunately, both of these ideas are operationalizable using the tools I develop in Chapter 3. Since entity extraction is a substantial part of the network-based measurement scheme I propose, a simple way to count the number of actors involved in the execution of each statute is to count the *number of unique entities* identified in each law. Though this measure does not focus on actors involved in "core regulatory functions" as in Farhang and Yaver (2016)'s measurement scheme, it still likely taps the same underlying concept as the more focused count these authors employ.

Measuring the frequency with which multiple actors are empowered to execute the same statutory function is more complicated. Unfortunately, many standard network statistics that relate to the underlying concept of interest are not comparable across networks with differing numbers of nodes. For example, consider a statistic like network density. For any given network, the density value associated with that network can be interpreted as the ratio of the number of ties actually observed compared with the total number of possible ties, given the number of nodes in the network. For bills with a large number of nodes, this statistic gives a reasonable estimate of the degree of overlap present in the bill's implementing network. However, for bills with a smaller number of nodes, this statistic is poorly-behaved. For an extreme example, consider a bill with only two nodes; because only one edge can potentially be observed, the density value for this bill can be zero or one. Because of this phenomenon, density values for small networks will be forced towards the extremes of the dependent variable's range, creating problems for both estimation and interpretation. Other plausible bounded statistics like the average clustering coefficient - which measures the the prevalence of non-overlapping "cliques" in a given network - are similarly problematic.

To avoid these issues, I instead measure the frequency of overlapping authority structures for a given bill using the *average degree* of that bill's implementation network.[5] In a weighted network, the degree of a given node is defined as the sum of all edge weights for edges connected to that node. The average degree of a network is therefore defined as the average degree value across all nodes in the network. Larger values on this statistic indicate a more densely connected implementing network, while smaller values indicate a more "siloed" implementing structure. Normalizing this statistic by the number of nodes in the network both eases comparison of this

[5]With the average degree defined to be zero for networks with zero named entities.

statistic across networks and avoids problems with overdispersion. Without normalization, the upper bound of this quantity scales with the squared number of nodes in the network, leading to extreme overdispersion and creating problems for estimation.

A list of the top twenty bills by each measure is given in Table 4.1. Overall, the average bill in my dataset possesses 9.5 nodes and an average degree of 5.3. However, both of these measures display substantial variation, and are characterized by both frequent zero values (27% by total nodes; 41% by average degree) and extreme overdispersion (maximum values of of 566 total nodes and an average degree of 708). As in Chapter 3, all of these values display substantial face validity; of the top twenty bills by the node count measure, nine are appropriations bills, seven are defense authorization bills, and the remaining four (the American Recovery and Reinvestment Act of 2009; two versions of the Food, Cnservation, and Energy Act of 2008; and the Energy Policy Act of 2005) are all are high-salience bills which involve substantial construction of new programs and disbursement of federal funds. Results are similar for the average degree measure, though some codification bills also display large fragmentation values by this measure.

Like Farhang and Yaver (2016) I find that both of my candidate measures are highly related. As shown in Figure 4.2, plotting the average fragmentation value for each measure and each session of Congress shows that these statistics essentially move in parallel throughout the period covered by my dataset. Because both measures are zero-inflated with substantial overdispersion, simple statistical measures of association like pairwise correlation values are misleading in this context; however,

Table 4.1: Top-twenty most fragmented bills, by two measures

Bill Title	Total Nodes	Average Degree
Omnibus Appropriations Act (2009)	566	411.2
Consolidated Appropriations Act (2008)	534	671.4
Consolidated Appropriations Act (2010)	458	707.6
Consolidated Appropriations Act (2012)	438	458.6
Defense Appropriations Act (2011)	448.6	
Consolidated Appropriations Act (2005)	306	277.7
Consolidated Appropriations Act (2004)	277	210.3
National Defense Authorization Act (2010)	270	43.1
National Defense Authorization Act (2015)	262	58.7
National Defense Authorization Act (2012)	245	40.2
Further Appropriations Act (2012)	238	380.1
Omnibus Appropriations Act (1990)	232	19.8
National Defense Authorization Act (1996)	229	35.8
Omnibus Appropriations Act (1999)	223	157.1
Food, Conservation, and Energy Act (2008)[1]	220	19.6
National Defense Authorization Act (1998)	214	30.6
American Recovery and Reinvestment Act (2009)	214	89.9
National Defense Authorization Act (2007)	212	38.5
Energy Policy Act (2005)	209	22.4
Food, Conservation, and Energy Act (2008)[1]	209	19.4

Bill titles shortened for visual purposes. Bills ordered based on total node value. Bolded titles are in the top twenty by both measures.

[1] A near-duplicate version of the Food, Conservation, and Energy Act (2008) were passed in 2008 to address drafting errors in the original version of the bill. In this table, the first entry for this bill refers to H.R. 2419, and the second refers to H.R. 6124.

transforming each variable by a log-plus-one transformation leads to a pairwise correlation value of 0.95. Because of this overlap, for the remainder of this chapter I focus on results generated using the node-count measure of my dependent variable.

4.3.2 Predictor Variables

As I describe in Chapter 2 and at the outset of this chapter, my key theoretical contention is that previously observed relationships between partisanship and downstream design of legislation should be moderated by policy area and by the public salience of the bill in question. To operationalize executive-legislative preference disagreements, I use a simple binary indicator, which consists of a dummy variable indicating whether the same party controlled both the executive and legislative branches at the time that the bill was passed. Based on existing theory, on average we should expect bills passed under unified government should contain less fragmented implementing structures than their counterparts passed under divided government. However, we should expect this relationship to be weaker for lower-salience bills and policy areas.

Besides this basic partisanship variable, I also include two individual-level and two bill-level covariates. For my individual-level variables, I include covariates corresponding to the DW-NOMINATE score of the proposing member and an indicator variable denoting whether the proposing member was a part of the chamber majority. Generally speaking, more conservative members tend to be more skeptical of a

Figure 4.2: Comparison of two measures of legislative fragmentation

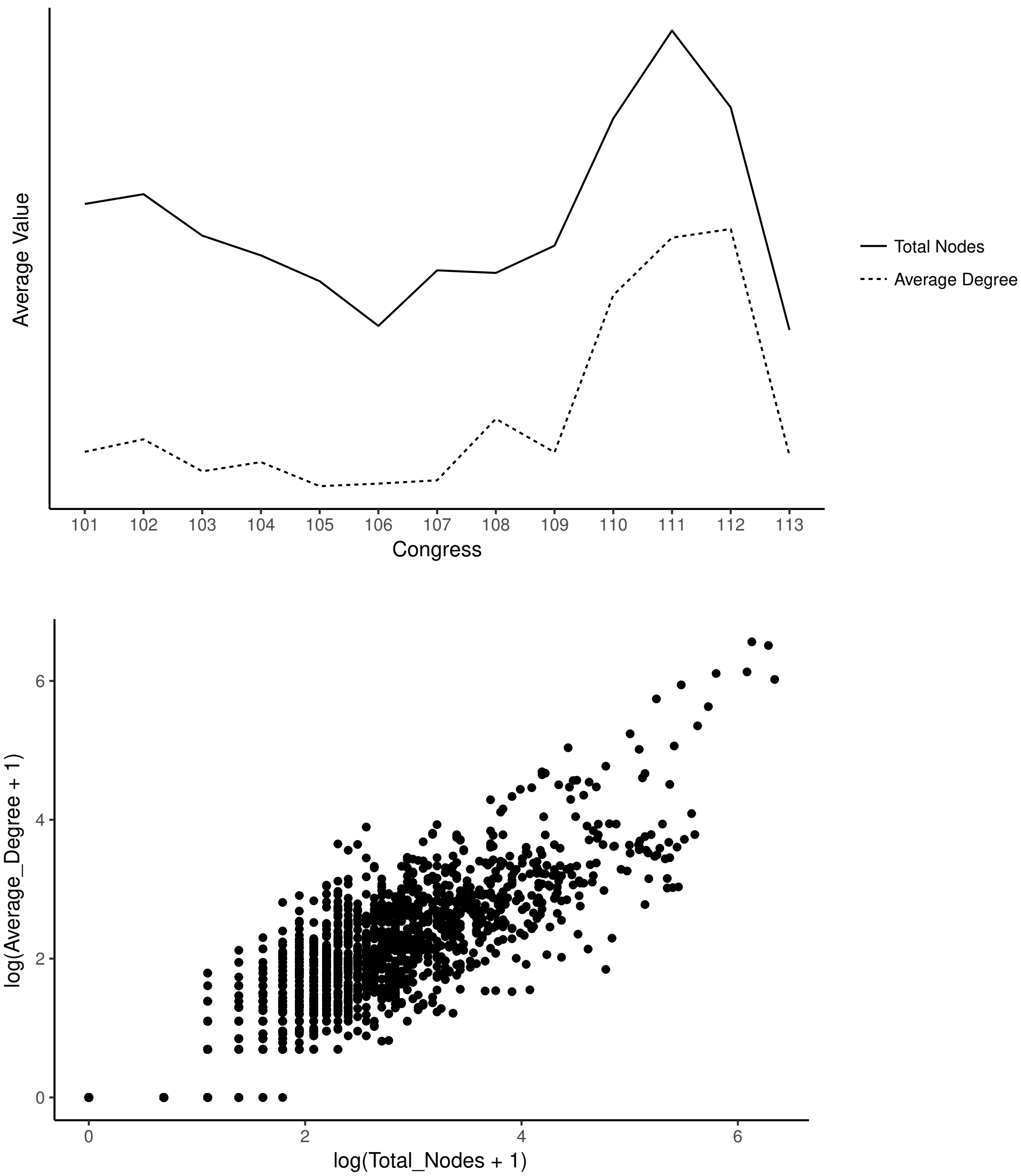


Table 4.2: Descriptive information for predictor variables

Variable	Mean	SD
Unified Government	0.31	0.46
DW-NOMINATE	0.1	0.5
Majority Sponsored	0.8	0.4
CQ Mention	0.3	0.5
$\sqrt{\text{Cosponsors}}$	2.6	3.0
Appropriation	0.03	0.18

Bill titles shortened for visual purposes. Bills ordered based on total node value. Bolded titles are in the top twenty by both measures.

strong administrative state; as a result, we might expect those members to be more willing to fragment implementing authority more frequently. By contrast, bills proposed by members of the chamber majority are likely to address higher-salience issue areas, making legislators more willing to craft complex and fragmented implementing structures.

For my bill-level variables, my primary theoretical quantity of interest is the *salience* of the bill in question. I operationalize this idea using two variables. First, to measure the broader public salience of a particular bill, I follow Volden and Wiseman (2014) and use a binary indicator denoting whether the bill in question was mentioned in the CQ Almanac's year-end summary of Congressional activity. Second, to measure salience within Congress I included a predictor corresponding to the square root of the number of cosponsors for each bill. Compared with my other variables, this measure is somewhat problematic; since important and time-sensitive bills often bypass the ordinary lawmaking process (Sinclair 2016), the authors of these bills may not have the time to gather a substantial number of cosponsors.[6] As a result, this measure may perform badly for appropriations bills, authorization bills, and other "must-pass" emergency legislation. As an additional control, I therefore include a dummy variable indicating whether the bill in question was an annual appropriations bill or a defense authorization bill, which should serve as a partial proxy for "must-pass" legislation of this kind.

Finally, to capture the interactive relationship between salience and legislative/executive preference disagreements I posit in Chapter 2, I include an interaction term between my `Unified_Government` and `CQ_Mention` variables. As I describe in Chapter 2, I expect the effects of preference disagreements on downstream design of legislation to be strongest on high-salience laws. When addressing high-salience policy issues, legislators should pass substantially simpler laws when operating under unified government than under divided government. By contrast, I expect the relationship between divided/unified government and legislative fragmentation to be attenuated for low-salience bills. Put together, these expectations imply that the interaction term between the `unified` government and `CQ_Mention` variables should be *negative* on average, though this relationship may vary some by policy area.

To operationalize policy area, I rely on the Congressional Bills Project's bill-level

[6]For example, contrast the Patient Protection and Affordable Care Act (Pub. L. 111-148) with the American Reinvestment and Recovery Act (Pub. L. 111-5). Though both bills were highly salient and the latter passed by substantially larger margins than the former, former bill had 40 cosponsors while the latter had only 9. This differential likely reflects the speed with which the bailout bill was enacted, compared with the more measured process for the Affordable Care Act.

policy codes. This variable follows the Comparative Agendas Project's coding scheme covering some 20 major topic codes. As shown in Figure 4.1, all major topic codes are represented in this period, though some topics are substantially more common than others.

4.4 Modeling

To model the node-count measure of fragmentation I propose, I employ a Bayesian hierarchical hurdle model, represented using the following likelihood function:

$$p(y_i|\rho_i, \theta_i, \phi) = \begin{cases} \rho_i & \text{if } \ y_i = 0 \\ (1-\rho_i)\frac{NB(y_i|\theta_i,\phi)}{1-NB_{CDF}(0|\theta_i,\phi)} & \text{if } \ y_i = 1 \end{cases}$$

With y_i the node count for the i^{th} observation, NB the negative binomial density, and NB_{CDF} the negative binomial CDF.[7] We can (very) loosely treat this model as a two-step regression, in which we first estimate a logistic regression to determine whether a given observation is zero or non-zero. For the set of observations selected into the non-zero component, we then estimate a negative binomial regression, in which the likelihood for the model is truncated at one.

[7] As defined using the location-scale negative binomial parameterization in Carpenter et al. (2016), in which $E(y_i) = \theta_i$ and $Var(y_i) = \theta_i + \frac{\theta_i^2}{\phi}$. Since ϕ is not subscripted, this structure implies a conditional constant variance assumption, in which I assume the variance of each observation to be constant for a given value of θ_i.

The "hurdle" component of this model refers to the mixture implied by the piecewise likelihood expression. Briefly, a hurdle model is an example of a mixture model, in which we treat the dependent variable as a mixture of two distinct probability distributions. In the context of this project, I expect to encounter two types of bills: a "standard" type, which increases, decreases, or otherwise modifies the jurisdiction of one or more governmental actors, and a "non-administrative" type, which does not alter the jurisdiction of any actor. Examples of the latter type include "commemorative" bills[8] or bills which consist of technical amendments, corrections, or updates to other pieces of legislation.[9] Laws of this kind are likely to follow a different data-generating process than other bills contained in my dataset, with very few (usually zero) nodes and very few (usually zero) edges connecting any nodes that are present. To separate bills of this kind from the other observations in my dataset, I therefore treat my dependent variable as a mixture of a Bernoulli and a negative binomial distribution.[10] Bills selected into the Bernoulli component of the mixture represent bills with no named entities or ties, and which are therefore unlikely to address the composition of the administrative state.

[8]E.g. Pub. L. 102-262, "A bill to designate the United States Courthouse located at 111 South Wolcott in Casper, Wyoming, as the 'Ewing T. Kerr United States Courthouse'." Recall that I filter bills identified as "commemorative" from the dataset before estimation; however, since the method used by the Congressional Bills Project to identify commemorative bills is heuristic and based on title keyword searches, some examples may slip through.

[9]E.g. Pub. L. 108-306, "To provide an additional temporary extension of programs under the Small Business Act and the Small Business Investment Act of 1958 through September 30, 2004, and for other purposes." This bill simply extends authorization for existing provisions of the Small Business Investment Act of 1958, and therefore provides no modifications to existing administrative jurisdiction.

[10]Since my average degree measure is (approximately) continuous and non-negative, I use a Bernoulli/gamma mixture for this dependent variable instead.

To incorporate covariates into this model, I define:

$$\rho_i = logit(X_i'\gamma_{z_i})$$
$$\theta_i = log(X_i'\beta_{z_i})$$

with X a $N \times (K+1)$ matrix of predictors as defined in Table 4.2, γ a $(K+1) \times M$ and β a $(K+1) \times M$ matrix of regression coefficients, and z_i an auxiliary $N \times 1$ vector of group assignments for each observation. This structure defines a standard hierarchical model, in which we separately estimate regression coefficients for both the logistic and negative binomial components of the mixture for each group. Here, I use the *policy area* of each bill as the group, yielding $M = 20$ unique groups for the dataset. For the X matrix, I include all predictor variables listed in §4.2.2 as well as an interaction between the `Unified_Government` and `CQ_Mention` variables in the count component of my model (yielding $K = 7$).

The theoretical motivation for this hierarchical coefficient structure follows directly from the expectations I outline in Chapter 2. Hierarchical Bayesian models are particularly useful when we expect the coefficients associated with most predictor variables to interact with some underlying group structure. For the purposes of this paper, I expect the relationship between most of my predictor variables and my dependent variable to vary according to the policy area. In the count component of the model, since the `DW-NOMINATE` variable is scaled from 1 (most conservative) to -1 (most liberal) I expect the coefficients on that variable to be more positive on issues prioritized by conservative lawmakers (e.g. defense) and negative on issues prioritized by their more liberal counterparts (e.g. civil rights). I further expect the coefficients on the intercept and the remaining predictor variables to be more positive for higher-visibility policy areas, which reflects the interaction between policy/issue salience and other predictor variables I outline in Chapter 2. For similar reasons, in the logistic component of the model I expect the intercept and the coefficients on the non-ideology predictor variables to be more negative for higher-salience policy areas.

Fitting this model with independent coefficients involves estimating $2M(K+1)$ regression coefficients, which may lead to high-variance coefficient estimates or problems with model convergence. To stabilize estimates, I therefore place a shared prior on each set of group-level coefficient estimates:

$$\mu_\gamma \sim MVN(\gamma, \Sigma_\gamma)$$
$$\mu_\beta \sim MVN(\beta, \Sigma_\beta)$$

With μ_γ and μ_β each a $(K+1)$ vector of top-level regression coefficients and Σ_γ and Σ_β each a $(K+1)\times(K+1)$ variance-covariance matrix. This prior structure "partially pools" coefficient estimates for each predictor, allowing the data to inform the model regarding the extent to which policy area-specific coefficients for each variable should be allowed to vary. In cases where we should expect a given coefficient's value to vary substantially across policy area, this prior structure allows coefficients to vary appropriately. However, in cases where the effect of a given coefficient is more uniform, the partial pooling structure prior structure allows coefficient posteriors to borrow precision from one another, reducing variance in these estimates.

To complete the model, I place a vague half-normal prior on the negative binomial scale parameter $\phi \sim N_T(0, 5)$. I then place priors on the variance-covariance matrices Σ_γ and Σ_β using the strategy suggested by Gelman et al. (2014). I first define auxiliary variables ν_β, ν_γ, Ω_β, and Ω_γ using the general form $\Sigma = diag(\nu)\Omega diag(\nu)$. This decomposition eases estimating by allowing me to place separate priors on the location and scale of the variance-covariance matrix for each set of coefficients. For numerical stability, I further decompose Ω using a Cholesky factorization such that $\Omega = LL'$, and place the following priors on the auxiliary variables:

$$\nu_\gamma, \nu_\beta \sim N_T(0, 10)$$
$$L_\gamma, L_\beta \sim LKJ(1)$$

With N_T a half-normal prior, and LKJ denoting the Lewandowski et al. (2009) correlation matrix distribution. $LKJ(1)$ reduces to an identity distribution over correlation matrices, which causes this prior to represent a flat prior over coefficient correlation. The prior on ν was selected to represent a vague but mildly informative prior, indicating a slight preference towards coefficient estimates that are smaller in absolute value. In most situations, priors of this kind aid numerical stability during estimation and improve posterior predictive performance.

To fit the model, I used the Stan programming language (Carpenter et al. 2016). I ran four chains, with 1000 warmup iterations and 3000 post-warmup iterations in each chain.[11] Visual plots suggested good mixing across chains, with $1 \leq \hat{R} \leq 1.01$ for all parameters and $n_{eff} \geq 1000$ for all parameters.[12] Posterior predictive checks (given in Appendix B.1) suggested that the model slightly under-predicted low ($y \leq 10$) values of the dependent variable; overall, however, model fit appeared acceptable.

4.5 Results

4.5.1 Hurdle Model

Top-level posterior means and credible intervals for the hurdle model coefficients are given in Figure 4.3. Estimated posterior values for all coefficients match the expectations I outline above; bills with a greater number of cosponsors and bills that receive a mention in the CQ Almanac are both significantly less likely to be selected into the zero component of the model. The partisanship and unified government covariates have posterior credible intervals that cross zero for both top- and lower-level estimates, suggesting that these coefficients are largely unrelated to the dependent variable in this component of the model.

Since the hurdle component of the model is essentially equivalent to a standard logistic regression, we can easily transform coefficient estimates to more substantively

[11] For further parameter details and diagnostics, see Appendix B.1

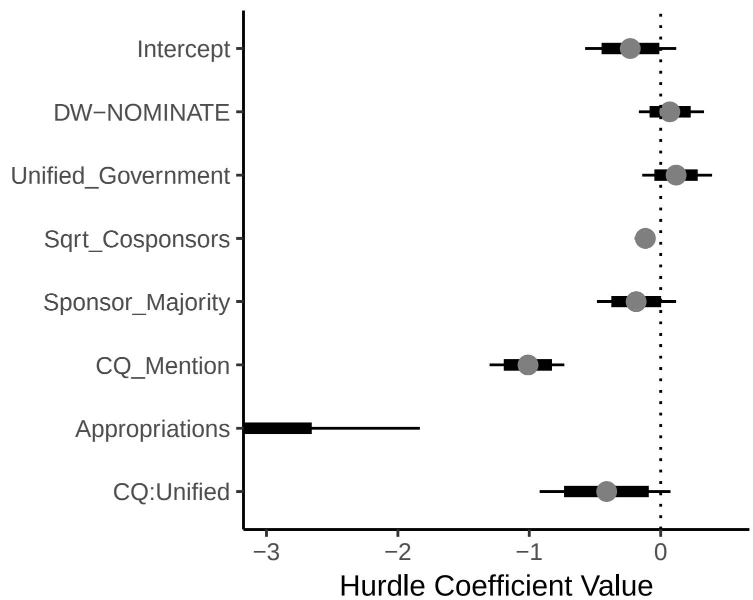
[12] With $\hat{R}$ a diagnostic quantifying the consistency of an ensemble of Markov chains, and n_{eff} a rough effective sample size calculation (Gelman et al. 2014).

meaningful values using standard techniques. For example, for bills passed under divided government, receiving a mention in the CQ year-end almanac reduces the odds that a given observation will have a value of zero on the dependent variable by 64%.[13] Policy-area specific estimates of this coefficient are roughly equivalent in in magnitude, suggesting that this variable's effect is approximately constant across policy areas. This finding is consistent with expectations; since most high-salience bills interact with some way with the administrative state, bills that receive press coverage are unlikely to be of the "non-administrative" type, no matter the policy area.

The coefficient on the cosponsor variable, by contrast, offers a good example of the payoff provided by the hierarchical coefficient structure I employ in this paper. As shown in Figure 4.3, the top-level posterior estimate of this coefficient is small but noticeably different from zero; a one-standard deviation decrease (≈ 3.0) in the `Sqrt_Cosponsors` produces an average 29% decrease in the odds that a given observation will have a value of zero on the dependent variable. However, as shown in Figure 4.4, estimates for this coefficient actually vary dramatically by policy area. For most bills, a greater number of cosponsors is associated with a small-to-moderate decrease in the probability that the dependent value will have a zero node count; however, for defense and transportation bills, a greater number of cosponsors actually *increases* the probability that a given bill will be of the "non-administrative" type. The reason for this difference is rooted in Congressional norms; since many of the

[13]Holding this coefficient at its posterior mean.

Figure 4.3: Top-level coefficients, hurdle component. The dependent variable is the probability that a given observation will have a fragmentation value of zero.



Dots indicate posterior mean values. Thick lines indicate 90% credible intervals, and thin lines indicate 95% credible intervals. Positive estimates indicate that an increase in the given parameter value makes the dependent variable more likely to take on a value of zero. Posterior mean and credible interval for the `Appropriations` coefficient are truncated for aesthetic purposes.

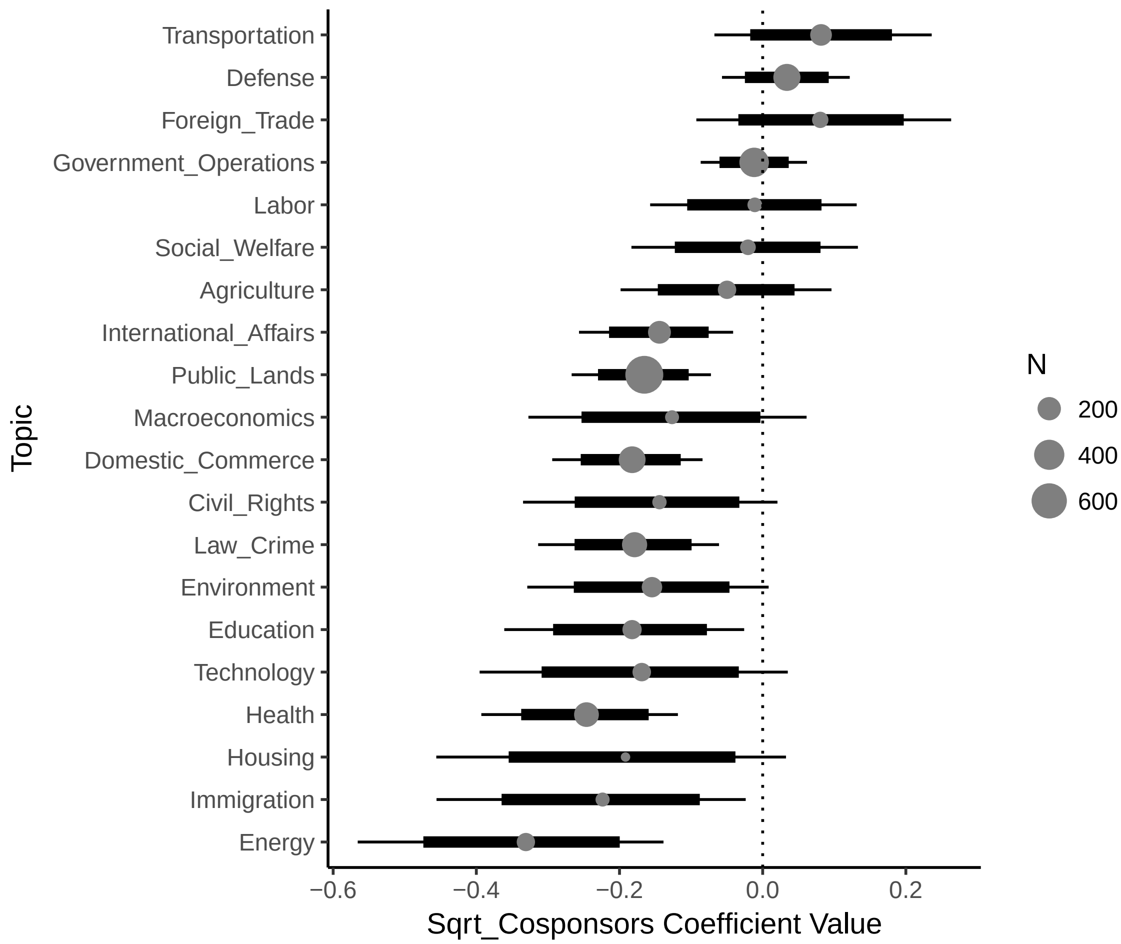
most fragmented defense and transportation bills are "must-pass" funding measures, these bills usually bypass normal procedures, and attract few or no cosponsors (Sinclair 2016). As a result, bills in these domains which do attract cosponsors are more likely to contain few or no named entities.

4.5.2 Count Model

Broadly, the count component of the model I present - which is restricted to bills that actually affect administrative authority - can be interpreted similarly to the hurdle component. Since the count component uses a log-link, we can interpret the exponentiated coefficient estimates as having a multiplicative effect on the expected value of the dependent variable. For example, since posterior mean coefficient estimate for the `Sponsor_Majority` variable is ≈ 0.36, exponentiating this estimate yields a predicted $\approx 44\%$ increase in fragmentation when comparing majority-sponsored bills to their minority-sponsored counterparts. Using a similar procedure for the `Appropriations` variable yields a predicted $\approx 320\%$ increase in fragmentation. Though enormous, this latter estimate is also sensible. As I describe earlier in this chapter, appropriations bills are some of the highest-salience and most contentious bills in my dataset, which should lead us to expect these bills to be unusually fragmented.

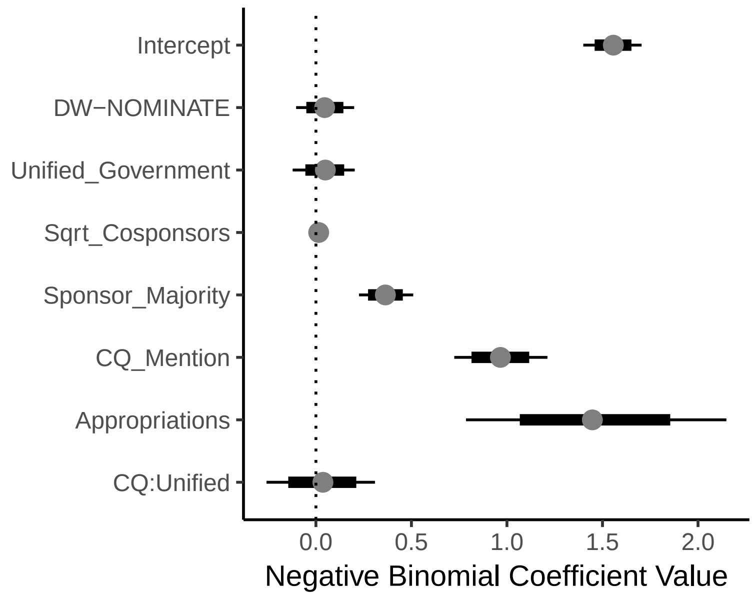
The coefficient on the `CQ_Mention` and `Unified_Government` variables partially match my expectations, but contain some unexpected results. As predicted, publicly salient bills are substantially more fragmented than their non-salient counterparts.

Figure 4.4: Second-level `sqrt_cosponsors` coefficients, hurdle component. The dependent variable is the probability that a given observation will have a fragmentation value of zero.



Dots indicate posterior mean values. Thick lines indicate 90% credible intervals, and thin lines indicate 95% credible intervals. Positive estimates indicate that an increase in the given parameter value makes the dependent variable more likely to take on a value of zero. Dot sizes scaled by the number of bills in each policy area.

Figure 4.5: Top-level coefficients, count model. The dependent variable is the number of nodes in a given bill, conditional on that bill's node count being non-zero.



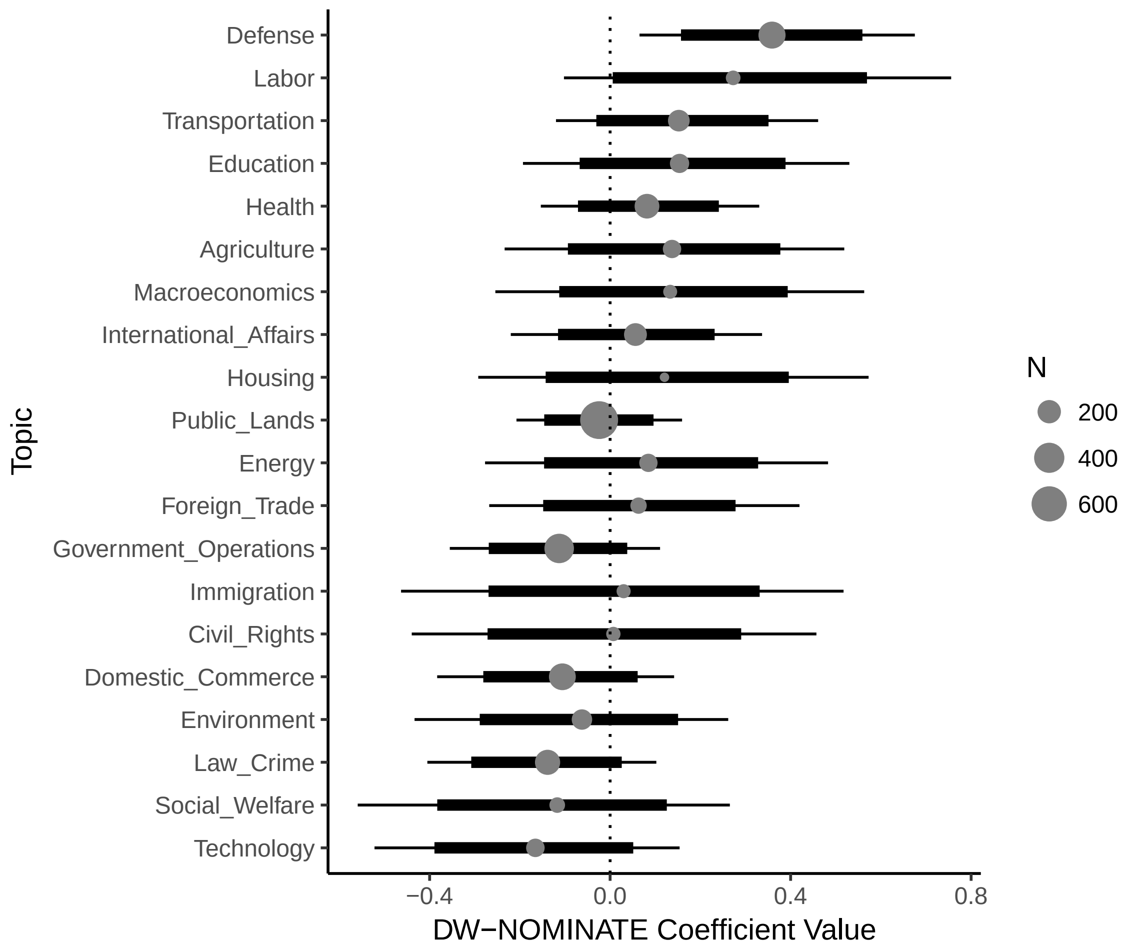
Dots indicate posterior mean values. Thick lines indicate 90% credible intervals, and thin lines indicate 95% credible intervals. Positive estimates indicate that an increase in a given coefficient increases fragmentation.

Contrary to my expectations, however, this relationship is essentially identical under divided and unified government. Averaged across policy areas, a mention in CQ's year-end almanac corresponds to a $\approx 164\%$ increase under divided government and a $\approx 172\%$ increase under unified government. Most surprisingly of all, bills passed under divided and unified government display essentially no differences in fragmentation levels.

However, as in the hurdle model, focusing on top-level coefficients can conceal substantial effect heterogeneity. For example, consider the `DW-NOMINATE` variable. Averaged across policy areas the ideological orientation of the proposing member has little impact on the a bill's administrative structure. However, this broad view conceals some potentially interesting policy area-specific differences. Since the `DW-NOMINATE` variable is scaled from -1 (most liberal) to 1 (most conservative), positive coefficients indicate that more conservative members tend to propose more fragmented bills in a given policy area, while negative coefficient estimates indicate that more liberal members tend to propose more fragmented bills in a given area. As Figure 4.6 shows, bills in nearly all policy areas display coefficient values near zero on this scale. However, bills addressing defense - and, to a lesser extent, labor and transportation - actually display positive coefficient values, suggesting that more conservative members of Congress tend to propose more complex laws in these policy areas. These effect sizes are not particularly large - for example, a one-standard deviation increase in the `DW-NOMINATE` variable (≈ 0.45) would be predicted to produce a $\approx 17\%$ increase in fragmentation for a defense bill - but they roughly track with the expectations I outline in Chapter 2. Intuitively, we should expect conservative lawmakers to be differentially motivated to create more complex institutional structures on bills addressing defense and other policy areas differentially prioritized by conservatives.

Effect heterogeneity also helps explain some of the surprising top-level results in the `CQ_Mention` and `Unified_Government` variables. Since I interact these two variables, we cannot consider their effects in isolation, which complicates interpretation. Fortunately, the Bayesian approach I use to estimate the model in this chapter enables a simple solution. Starting with the `CQ_Mention` variable, to generate the estimated

Figure 4.6: Lower-level `DW-NOMINATE` coefficients, count model. The dependent variable is the number of nodes in a given bill, conditional on that bill's node count being non-zero.



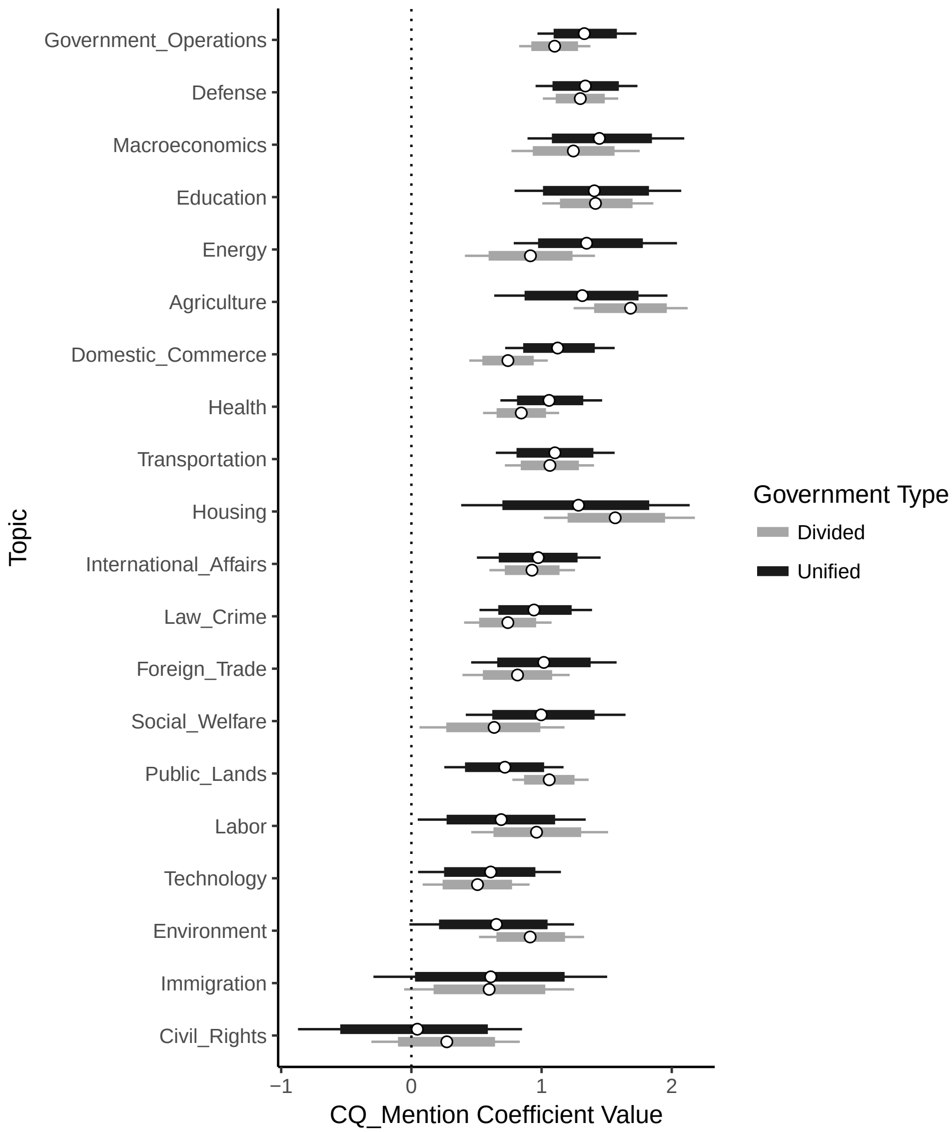
Dots indicate posterior mean values. Thick lines indicate 90% credible intervals, and thin lines indicate 95% credible intervals. Positive estimates indicate that an increase in the coefficient increases the bill's estimated fragmentation value. Dot sizes scaled by the number of bills in each policy area.

coefficient when `Unified_Government` = 1, we can simply add the posterior draws from each iteration for the `CQ_Mention` and `CQ_Mention:Unified` interaction variables, and use the results to produce posterior mean and credible intervals for this scenario. To generate estimates when `Unified_Government` = 0, we can simply use the raw posterior draws for the `CQ_Mention` coefficient.

The results of this procedure are shown in 4.7. As shown in Figure 4.7, though this variable's estimated effect is large and positive in for most policy areas, the scale of its effect varies dramatically. Under unified government, being mentioned in CQ's year-end almanac yields a predicted 300-400% increase in fragmentation in policy areas like government operations, defense, and macroeconomics. By contrast, public salience affects bills addressing civil rights, immigration, and environment much more modestly. This broad pattern remains similar under divided government, though the average effect size is larger and rankings across policy areas are somewhat shifted.

We can use a similar strategy to investigate effect heterogeneity in the `Unified_Government` variable. As I predict in Chapter 2, visual inspection of Figure 4.8 suggests that the relationship between partisanship and downstream design of legislation is attenuated for non-salient bills. In other words, bills that receive little or no public attention possess essentially equivalent implementing structures when passed under unified or divided government. By contrast, bills that receive substantial public attention vary more noticeably when passed under divided and unified government, with most coefficient estimates being larger in absolute value for salient bills. A simple Bayesian posterior probability calculation verifies this visual inspec-

Figure 4.7: Lower-level `CQ_Mention` coefficients, count model. The dependent variable is the number of nodes in a given bill, conditional on that bill's node count being non-zero.

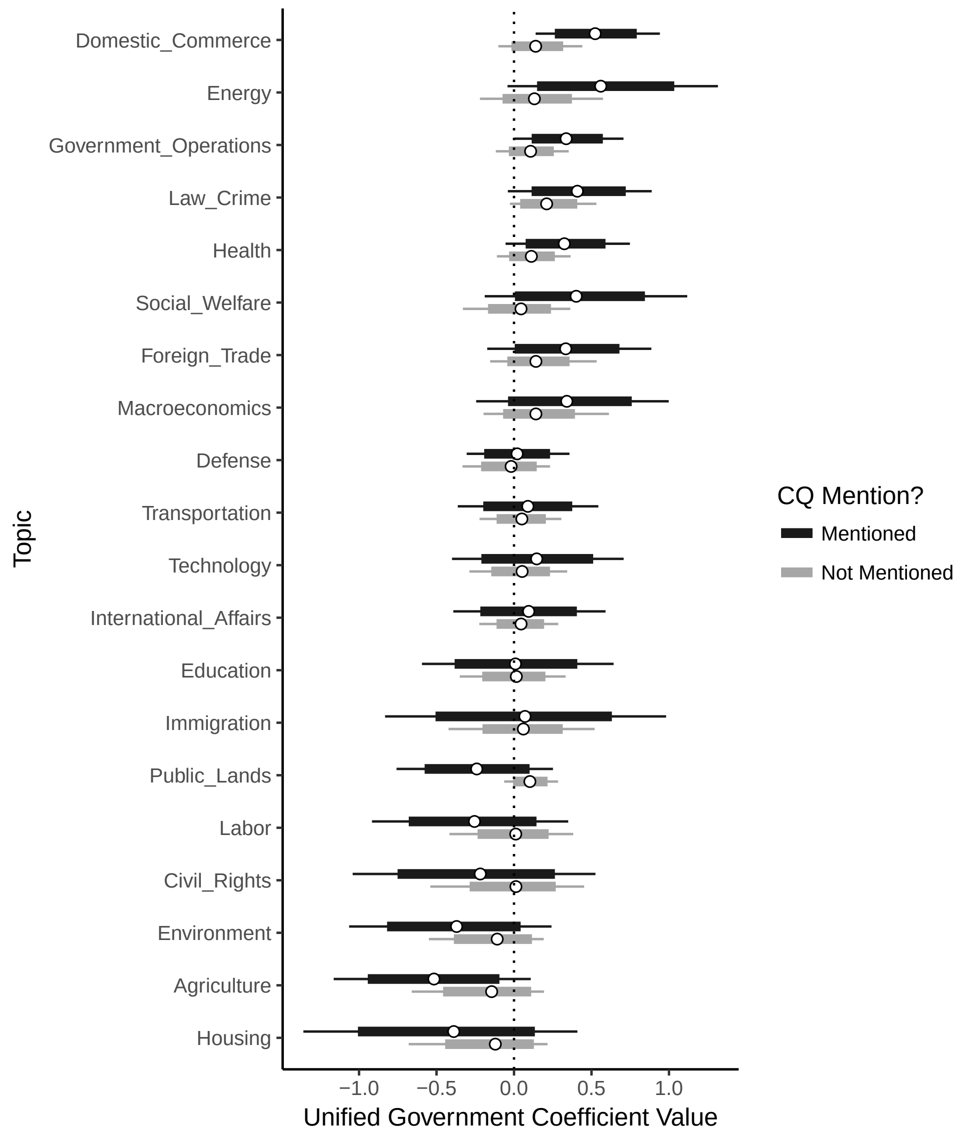


Dots indicate posterior mean values. Thick lines indicate 90% credible intervals, and thin lines indicate 95% credible intervals. Positive estimates indicate that an increase in the coefficient increases the bill's estimated fragmentation value.

tion, suggesting that - averaged across all policy areas - there is an estimated 78% posterior probability that a given coefficient estimate will be larger in absolute value for high-salience bills than low-salience bills. This attenuation pattern is largest for policy areas like domestic commerce, energy, government operations, and agriculture, but is noticeable in most policy areas in the dataset.

These findings help explain the surprising top-level results I present in Figure 4.5, and offer partial support for the hypotheses I offer at the outset of this chapter. However, the overall weakness of the relationship between divided and unified government and downstream allocation of authority is surprising, and merits further investigation. Interestingly, the results I present here actually align with those from at least one other major study. In their book, Huber and Shipan (2002) limit their attention to state-level health policy bills, and find a positive relationship between discretion (an idea related to fragmentation) and unified government. This finding matches the coefficient I present for health bills, and offers some reassurance that coefficients I report are accurate. One possible explanation for the difference between bills addressing the policy areas with positive coefficients on the `Unified_Government` variable and the remainder of the dataset relates to the type of bills passed in these areas. Since bills in health, government operations, and domestic commerce all frequently involve substantial appropriations, lawmakers acting under unified government may be more willing to fragment authority on these bills in order to protect their favored policy programs. Under divided government, by contrast, appropriations bills may create fewer new programs, creating fewer incentives for lawmakers to create com-

Figure 4.8: Lower-level `Unified_Government` coefficients, count model. The dependent variable is the number of nodes in a given bill, conditional on that bill's node count being non-zero.



Dots indicate posterior mean values. Thick lines indicate 90% credible intervals, and thin lines indicate 95% credible intervals. Positive estimates indicate that an increase in the coefficient increases the bill's estimated fragmentation value.

plex implementing structures for these bills. Clearly, however, probing this and other possible explanations further represents a direction for future research.

Two other possible explanations for this divergent finding are *time* and *selection*. Since Epstein and O'Halloran (1999) and Farhang and Yaver (2016) both study legislation over the period from 1945 onwards, the strong divided government effect they identify may actually be restricted to the earlier period in their dataset. However, this explanation seems unlikely. If a divided/unified government effect is present, increasing Congressional polarization in recent decades should *increase* rather than *decrease* the size of this coefficient, leading me to observe a larger coefficient on this variable than that observed in other studies.

A more likely explanation for the divergence between my findings and those of most previous studies is a *selection* problem. Because Epstein and O'Halloran (1999) and Farhang and Yaver (2016) only examine laws drawn from Mayhew (1991)'s list of "significant" legislation, their studies necessarily select out all but the highest-attention bills in the dataset. However, based on the expectations I present in Chapter 2, this set of bills is precisely the set in which we should expect to see the strongest relationship between divided government and fragmentation of authority. When addressing high-salience policy areas, politicians are particularly incentivized to design administrative structures carefully and to be suspicious of executive malfeasance. However, since bills that reach this level of public salience are relatively rare, in most cases considerations like policy area and partisan policy priorities are more influential on the implementing structure in a given bill. Re-estimating the model I present to only include high-salience bills might help to probe this explanation. However, a narrower measure of public salience than the one I employ might be necessary in order to observe this effect.

4.6 Conclusion

Overall, the results I present provide support for many of the hypotheses I present in Chapter 2 and elsewhere in this dissertation. Throughout the models I present, the *salience* of a given bill (as measured by mention in CQ's year-end almanac and the number of cosponsors that bill attracts) consistently represent some of the most important predictors in the model. High-salience bills are dramatically more likely both to affect administrative jurisdiction and to contain more fragmented implementing structures.

Also as predicted, many of the relationships I examine are conditional on *policy area*. In the hurdle component of the model, for emample, increasing cosponsor counts decrease the probability that a bill will be in the "non-administrative" category for most policy areas, but actually increase this probability for defense bills and other bills in policy areas with many measures passed outside the usual legislative process. In the count component of the model, the unified government coefficient is also contingent on policy area. Averaged across policy areas, unified-government bills are no more or less fragmented their counterparts passed under divided government. But, for politically salient bills, the scale and direction of this relationship actually varies substantially across policy areas, with bills in some policy areas displaying a negative relationship between unified government status and fragmentation and some displaying a positive relationship.

Though encouraging, these results leave space for improvement. In particular, *bills* may not represent the proper unit of analysis for this study. As I show in Table 4.1, the highest-complexity bills in my dataset largely consist of expansive omnibus laws. Since omnibus bill components are usually proposed individually, a more reasonable approach may be to break these omnibus laws into their constituent components, and consider each bill individually. This approach also allows me to more precisely identify the sponsoring member for each bill, which offers additional measurement advantages.

These limitations aside, the results I present emphasize the importance of both the theoretical ideas and the measurement techniques I introduce in this dissertation. In existing work on legislative fragmentation, measurement constraints have forced authors to restrict their attention to single policy areas or to "significant" policy areas. However, as I show, both of these factors substantially affect the design of legislation, both on their own and by structuring the relationships of other predictors. Without the measurement techniques I develop in this dissertation, these findings would not have been possible to produce, emphasizing the importance of scalable, broadly applicable measurement techniques for applied work.

Chapter 5

Information Extraction in the U.S. Code

In Chapter 4, I applied the measurement techniques I developed in Chapter 3 to analyze allocation of authority in enacted American legislation. As I argue throughout this dissertation, constraints imposed by existing measurement techniques have forced researchers interested in formal allocation of authority to limit their scope, usually to a subset of "significant" laws or to single policy areas. By contrast, the text-based approach I propose is more scalable, and allows researchers to study allocation of authority patterns in a broader set of texts.

However, like all empirical studies, the results I present in Chapter 4 contain shortcomings. In particular, by using statutes as my unit of analysis, I potentially ignore the effects of *policy context*. Except in rare cases, most bills do not seek to solve an entirely new policy problem. Instead, bills usually operate within an existing legal setting, which contains preexisting legal rules and an existing decision-making structure. Treating bills as independent observations ignores this structure, creating potential problems for inference.

As a partial solution to this problem, in this chapter I turn to the United States Consolidated Code. The Code is an official compendium of all enacted federal-level American legislation, organized (roughly) by policy area and updated for each new enacted law. Usefully for this project, the Code is also *versioned*. Each year, the Office of Law Revision Counsel releases a new edition of the Code, which incorporates all additions, deletions, and modifications to the Code generated through legislation enacted in the previous year. As a result, by comparing year-to-year versions of the Code, rather than bill-level enactments, I can implicitly control for policy context. Supplementing the bill-based analysis I provide in Chapter 4 with the Code-oriented analysis I offer in this chapter therefore offers a useful robustness check, as well as demonstrating the versatility of the methods I develop in this dissertation.

The remainder of this chapter proceeds in three parts. First, I provide a more detailed overview of the (potential) problems caused by the bill-based unit of analysis I use in Chapter 4. I then introduce the US Code dataset I use for the remainder of this chapter, and describe both its desirable features and its limitations. Finally, using the methods I develop in Chapter 3, I analyze changes in fragmentation patterns across the 1994-2016 versions of the US Code. After controlling for policy area and lagged fragmentation, I find that fragmentation increases slightly faster under unified than divided government. However, this is only substantively significant in a few policy areas, which mirrors the non-effect I observe in the bills dataset in Chapter 4.

5.1 From Statutes to the Code

5.1.1 The Problem with Context

Like most studies of formal allocation of authority, I use *bills* as my unit of analysis in Chapter 4. This choice is a natural one, and offers both theoretical and data-related advantages. From a theoretical perspective, most models of legislative behavior view bills as a fundamental organizing unit. Standard ideal-point models, for example, use votes on bills to identify the ideological orientation of members of Congress (e.g. Poole and Rosenthal 2000; Clinton et al. 2004), and agenda-setting studies use bills as a measure of Congress's policy agenda (Baumgartner et al. 2009). Most formal modeling studies designed to examine legislative/executive relationships also use bills as their unit of analysis; for example, Epstein and O'Halloran (1999), Huber and Shipan (2002), Volden (2002), and Farhang and Yaver (2016) all use bills as the basic bargaining unit in their models of legislative/executive interactions. Put together, studies like these offer a strong set of baseline expectations, variables, and measures for all stages of the bill-writing and negotiation processes, offering a strong theoretical foundation from which to work.

Bill-level data resources are similarly well-developed. For most key bill-level concepts - such as issue salience, policy area, and party positioning - scholars have produced agreed-upon operationalizations, which offer a natural starting point for empirical work. For example, in Chapter 4 I use the Congressional Bills Project's policy area codes, Volden and Wiseman (2014)'s bill salience data, and congress.gov's bill cosponsorship metadata. Datasets like Stewart and Woon (2011)'s committee membership information also make it easy for scholars to connect bills with the characteristics of their sponsors, which I use to identify the party and ideological orientation of the proposing member of each bill. Datasets like these enable researchers to easily collect detailed bill-level covariate sets, enabling rich, flexible models like those I fit in Chapter 4.

Despite these advantages, bills also possess some serious flaws as a unit of analysis. As I note above, these challenges can be broadly summarized as a problem of *policy context.* Generally speaking, legislators enact legislation to address some already-known issue, either by updating existing rules and guidelines or by reconfiguring decision-making structures. As a result, the characteristics of any given bill depend both upon contemporaneous factors - for example, the salience of the issue or policy area in question - and upon the characteristics of previous bills that address similar policy issues.

As a concrete example of this problem, consider the Patient Protection and Affordable Care Act (ACA). During my period of study, the ACA is one of the longest, most complex, and most fragmented bills I examine. These characteristics are undoubtedly due at least in part to factors present at the time the ACA was passed. During his presidential campaign, healthcare reform was one of Barack Obama's signature campaign pledges, making the issue unusually salient. Moreover, health policy more generally is a consistently salient policy area for most voters, ranking highly among non-economic policy areas in Most Important Problem-style surveys. However, at least part of the ACA's structure is also likely attributable to the *legacy* of previous healthcare legislation. Like most bills, the ACA is situated within an existing political, legal, and bureaucratic context. As a result, the complexity of its structure is at least partially a function of the existing complexity of the healthcare regulatory state.

Unfortunately, by treating bills as independent observations, the analysis I present in Chapter 4 cannot control for these kinds of contextual effects. Worse, it is difficult to see how a bill-centric study could even begin to address this shortcoming. Identifying the proper policy "legacy" of a particular bill would involve locating all previous legislation that addressed the same policy problem, in all or in part. This task is daunting enough for a legal expert studying a single bill, and is virtually impossible for an entire dataset.

5.1.2 A (Partial) Solution: the Consolidated Code

To address this problem, in this chapter I shift my unit of analysis from statutes to the US Consolidated Code. Briefly, the Consolidated Code is a compendium of all in-force American federal legal language organized by subject, which is maintained by the nonpartisan Office of Law Revision Counsel (OLRC). Unlike statute texts - which are written and organized according to political demands - the Code is primarily designed as a *legal research tool.* For most legal research applications, searching through individual statute texts is inconvenient. In their day-to-day work, lawyers are generally most interested in understanding the state of the law as it pertains to some particular policy issue or administrative rule. However, except in rare cases, the full state of the law on a single policy area is not contained in any one bill. Rather, regulations pertaining to a given issue are scattered across many different laws, each of which offers small modifications, additions, or subtractions to the legal rules in the area of interest.

The Consolidated Code is designed to address these shortcomings. Rather than being organized by date or by bill number, the Code is organized by subject. As a result, the text of any particular statute is rarely contiguous within the Code. Instead, the text of a particular statute is often divided between many different sections of the Code, in order to respect the Code's subject-specific organizational structure. The Code also reflects amending and repealing actions; so, if a particular bill alters or removes an existing piece of law, that change is also incorporated into the Code. In sum, the Consolidated Code therefore represents the *state of the law* at a given point in time, including any and all additions, deletions, and amendments introduced by an enacted bill up to a particular point in time.

Usefully for this project, the code is also *versioned*. Every six years, the OLRC publishes an updated version of the Code, which reflects all changes made during the interim period. Since 1994, the OLRC has also published online-only yearly versions of the Code, which provide additional year-on-year updates to the Code's content. These yearly versions, in particular, offer a natural way to address some of the problems created by using bills as a unit of analysis. Given the current year's Code version and the previous year's Code version, we can be confident that any *differences* between those two versions are attributable to political, legal, and policy factors specific to the last year, rather than to the legacy of previous legislative enactments. In other words, by focusing on the differences between consecutive versions of the Code, we automatically control for *policy context*. As a result, if (for example) we observe that the increase in fragmentation of the Code during periods of divided government is larger than the increase under unified government, we can be relatively confident that this effect is actually attributable to political factors during that period, rather than worrying that this pattern reflects a legacy of older policy choices.

No analytical approach is perfect, and the Code-based approach I propose in this chapter is no exception. Most notably, though the Code itself is attractive as an object of study, supplemental data resources for the Code are relatively limited. Most notably, no existing study connects individual sections of the Code to the statutes from which they were drawn. As a result, at present I cannot connect individual sections of the Code with bill- or individual-level metadata of the sort I rely upon to construct my models in Chapter 4. In addition, in certain cases, some modifications to the Code may not be attributable to political factors. Since 1994, the OLRC has undertaken six editorial reclassification projects, in which the OLRC unilaterally reorganized particular titles of the Code in order too provide a more useful subject-specific organization scheme. In addition, during my period of study various titles of the Code have also been enacted by positive law (re)codification legislation. Positive law codification is a process by which Code provisions organized by the OLRC are formally enacted into law by Congress. At least in recent cases, these positive law codification bills - which are themselves usually drafted by the OLRC -

also frequently involve substantial editorial reorganization of the Code Fortunately, as I describe in §5.2.2, these projects are well-documented[1], enabling me to identify major reclassification efforts relatively easily.

Put together, I suggest that these factors suggest that a Code-based analysis can act as a useful supplement to the statute-based analysis I present in Chapter 4. Unlike the statute-based study I present in Chapter 4, by focusing on year-to-year changes in the Code I can be confident that any changes observed are attributable to the political, legal, and policy factors present during the period in question. Unfortunately, since section-specific metadata are not available for the Code, we cannot be certain *which* year-specific factors caused any observed changes. Still, by comparing changes in the Code during periods of interest (for example, under unified versus divided government), we can gain some additional insight into the lawmaking process above and beyond that which would be possible by looking at bills alone.

5.2 Assembling the Consolidated Code Dataset

5.2.1 Parsing the Code

As I describe in the previous section, the US Consolidated Code is maintained by the Office of Law Revision Counsel (OLRC). Since 1994, the OLRC has issued yearly releases of the Code online in XML format, which contain all new legislative text,

[1] See the OLRC website for details.

repeal actions, and amendments to the Code contained in legislation passed in a given year. At time of writing, code versions for the years between 1994-2016 were available; as a result, my dataset for this chapter consists of the 23 years covered by this time period.

To parse the code, I follow a similar procedure to the one I outline in Chapter 4. First, I parse the Code according to its organizational headers. Usefully - and in sharp contrast to the enacted bill data I use in Chapter 4 - the OLRC's Code releases follow a highly consistent format and contain a substantial quantity of embedded metadata. As shown in Figure 5.1, the OLRC's code releases clearly and consistently denote section headers and legally operative language using a standardized HTML stylesheet. This information makes parsing and cleaning the Code straightforward. In particular, for each text I identify all tags with `class = section-head`, and assign all language with `class = statutory-body` to the preceding section header.[2] I then discard all editorial notes, citation information, and other extraneous text, and save the resulting output to a series of JSON files.

Second, I segment the Code into chapters. Unlike in Chapter 4 - where bills were my fundamental unit of analysis - there is no obvious fundamental "unit" of the Code. In principle, I could construct implementing networks for each title, chapter, subchapter, or any other organizational unit of interest. However, to maintain comparability with Chapter 4, I use *chapters* as my unit of analysis in this chapter. I make this

[2] Using the BeautifulSoup parser in Python 2.7.
[3] 1 U.S.C 101.

Figure 5.1: Example Consolidated Code embedded HTML metadata[3]

```
<h3 class="section-head">
  &sect;101. Commencement of term of office
</h3>

<p class="statutory-body">
  The term of four years for which a President and Vice
  President shall be elected, shall, in all cases,
  commence on the 20th day of January next succeeding
  the day on which the votes of the electors have been
  given.
</p>
```

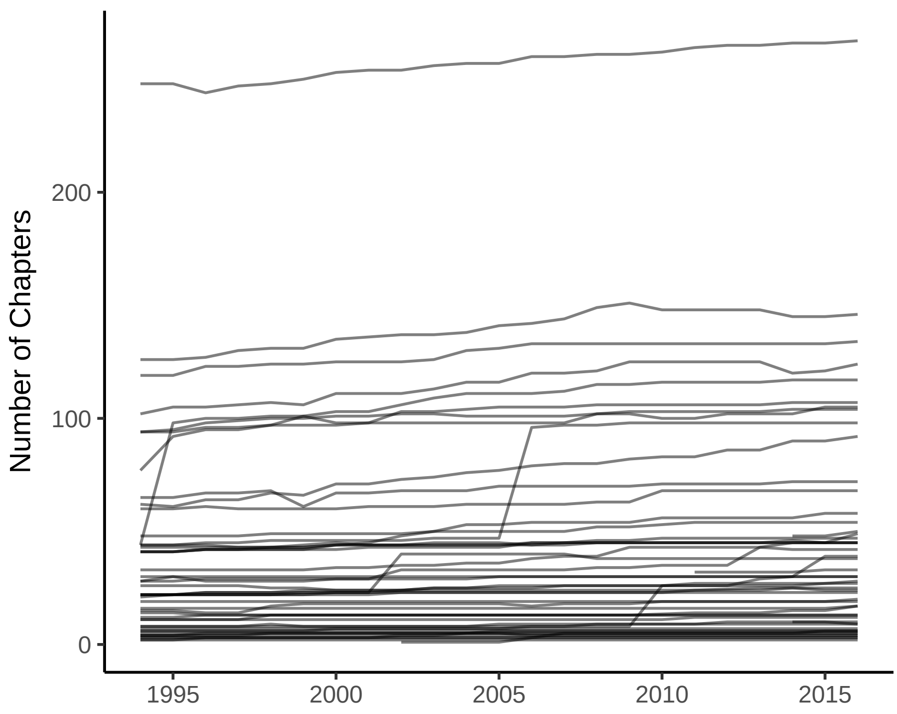
choice for two reasons. First, from an impressionistic standpoint, chapters in the US Code appear to be the organizational unit that most closely corresponds to the magnitude and substantive scope of the average bill. Titles, for the most part, are much longer and more thematically diverse than all but the largest bills, while subchapters and other lower-level sections are much shorter than the average bill. Second, from a practical standpoint, chapters represent the lowest-level organizational unit that is both present in every title and strictly higher in the organizational hierarchy than sections. Since I use sections to construct the ties in the implementing networks I generate, chapters therefore represent the most granular available organizational unit.

Some basic descriptive information regarding the final Consolidated Code dataset is shown in Figures 5.2 and 5.3. Overall, the dataset contains some 1,895 unique chapter-title pairs and 35,362 chapter-year observations, though not all chapter-title pairs are present in all years. By far the largest title (as measured by number of chapters) is Title 10, which covers the Department of Defense and the US Armed Forces. Title 42 - which covers public health, social welfare, and civil rights - is the next-largest, followed by Titles 18 (criminal law) and 16 (conservation and public lands). Averaged across the period, the average title has some 47 chapters, with a maximum of 267 (Title 10, 2016) and a minimum of one (various) titles and years. Unsurprisingly, the total number of chapters remaining roughly constant for most titles throughout the period I examine. Since most sections of the code are not modified in any particular session of Congress, we should expect chapter creation or modification to be a relatively rare event when averaged across the entire Code.

5.2.2 Extracting Implementing Networks

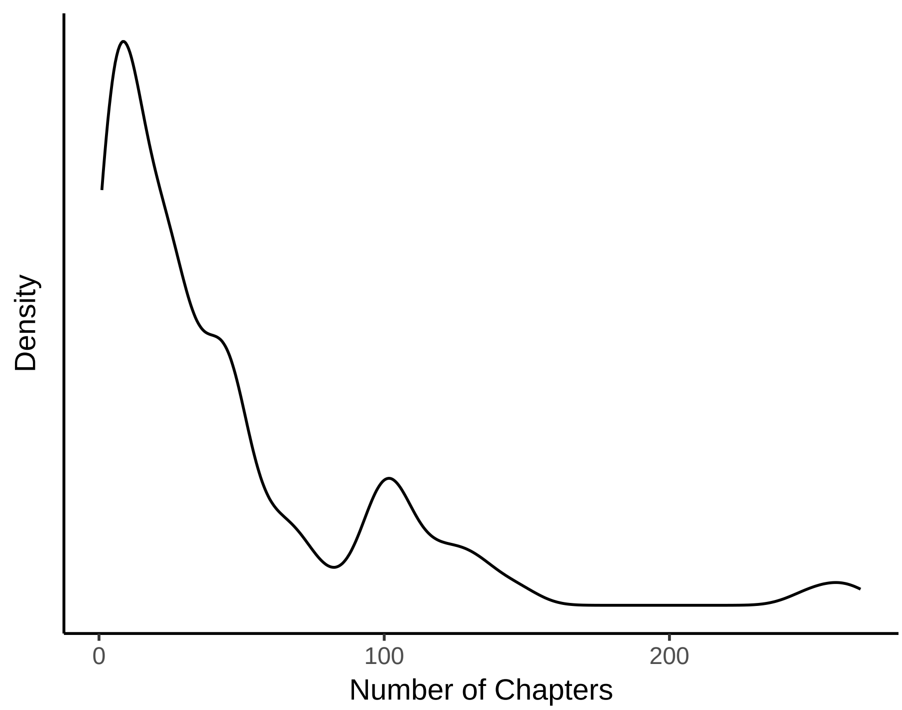
To construct chapter-level implementing networks, I follow a very similar procedure to the one I describe in Chapter 4. Beginning with the initial 1994 Code release point, I use the pre-trained neural network model to extract all named entities from each chapter's text. If two named entities co-occur within a particular section, I draw an edge between them, with the total edge weight in each case corresponding to the total number of instances in which two named entities co-occurred within a section. I then repeat this process for all chapters in the 1994 release point. For subsequent release points, I first check if each chapter's text has been updated across release points. If a given chapters text has been changed between release points, I re-run my analysis procedure on the new text, and save the output. If not, I simply carry the

Figure 5.2: Chapter count, by US Code title



Lines indicate a single US Code title. Chapter counts observed yearly, using the scraped and segmented US Code dataset.

Figure 5.3: Chapter count frequency, by year and US Code title



Unit of analysis is title-year. Chapter counts generated using the scraped and segmented US Code dataset.

previous release point's implementing network forward to the current one, allowing me to avoid expending computing time re-analyzing identical chunks of text.

As an example of this procedure, in Figures 5.4 and 5.5 I provide outputs for 12 U.S.C. 29 from the 1994 and 2016 editions of the US Code. Briefly, this chapter contains language originally enacted as part of the Home Mortgage Disclosures Act of 1974, which requires financial institutions to provide certain mortgage data to federal regulators. Subsequent legislation dvided responsibility over data management between the Comptroller of the Currency, the Office of Thrift Supervision, the Federal Deposit Insurance Corporation, National Credit Union Administration Board, and the Secretary of Housing and Urban Development, depending on the type of financial institution. Reassuringly, all of these actors are placed in central positions in the network visualization shown in 5.4, with more minor Congressional actors and other institutions placed at the outside of the plot. Examining network centrality statistics reinforces this impressionistic view, with the Federal Deposit Insurance Corporation (eigenvector centrality of 0.46) and the Office of Thrift Supervision (0.43) occupying the most central positions.

Comparing the 1994 version of this chapter to the 2016 edition shown in Figure 5.5 shows a broadly similar implementing structure. This invariance is sensible; since the Home Mortgage Disclosures Act of 1974 has not been heavily modified since its enactment, we should not expect to see substantial changes in its structure. Among the changes from 1994 to 2016, the most prominent differences are the absence of the Office of Thrift Supervision, and the addition of the Bureau of Consumer Financial

Figure 5.4: Implementing network, 12 U.S.C. 29 (1994)

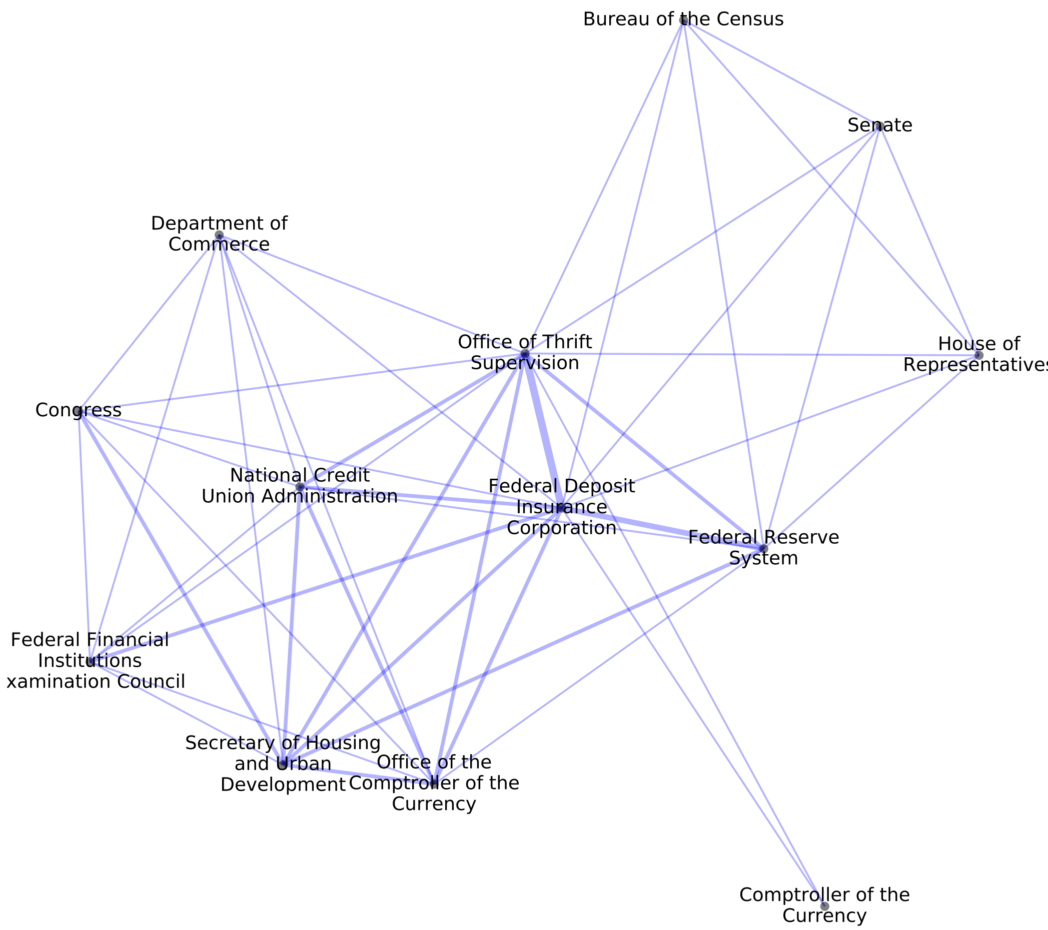
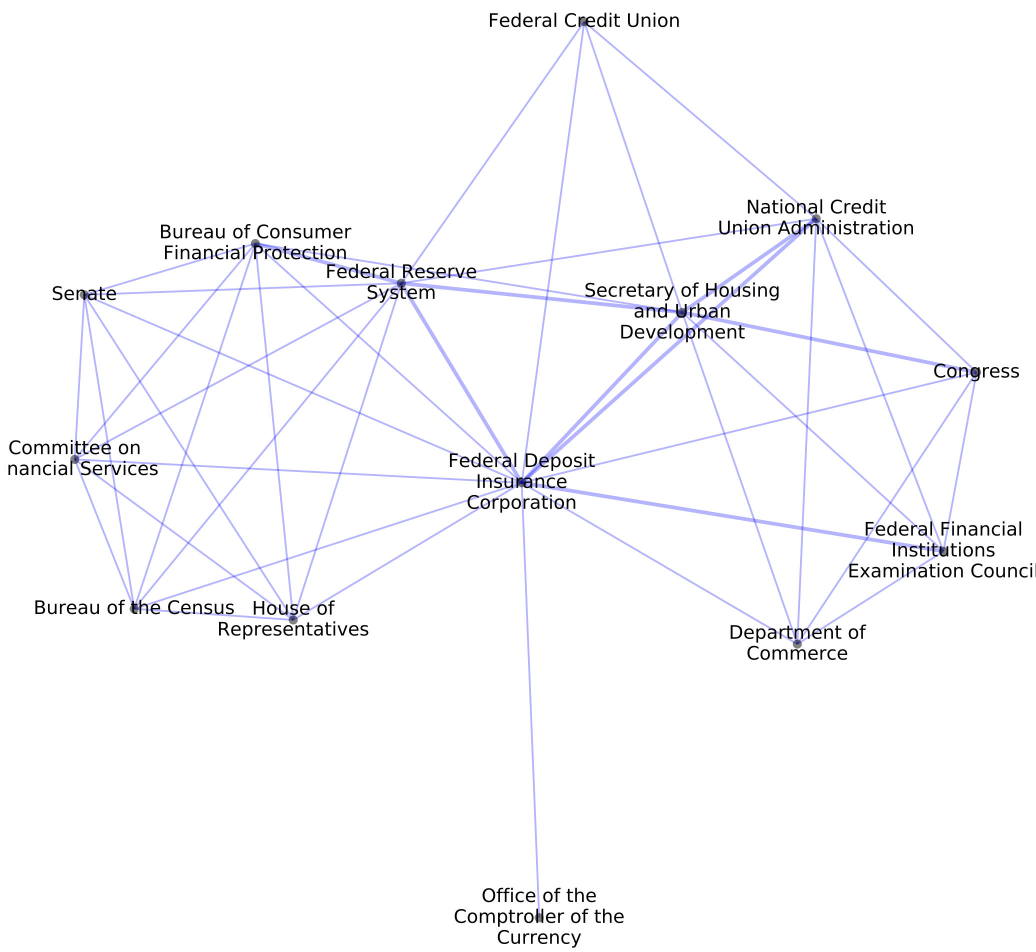


Figure 5.5: Implementing network, 12 U.S.C. 29 (2016)



Protection. As before, these changes reflect the regulatory history of this policy area. Following the 2007-2008 Financial Crisis, the Office of Thrift Supervision was one of the main institutions implicated in some of the regulatory failures that contributed to the crisis events. Because of these shortcomings, the Dodd-Frank Act[4] dissolved the Office of Thrift Supervision, and redistributed its responsibilities redistributed to the newly-created Bureau of Consumer Financial Protection and related financial regulation agencies.

5.2.3 Identifying Non-Substantive Code Revisions

One challenge unique to the Consolidated Code is problem of separating non-substantive Code revisions from substantive changes to the law. Unlike ordinary legislation, the US Code is organized and managed by the Office of Law Revision Counsel (OLRC). As part of its duties, the OLRC periodically conducts non-substantive updates to the structure of the Code, in which sections of the Code are relocated, split, combined, or (for obsolete sections) deleted. As I describe in §5.1.2, these efforts can be (roughly) divided into two types: *editorial reclassification* - which is conducted unilaterally by the OLRC - and *positive law codification* - which is officially authorized by enacting legislation passed through the normal legislative process. Non-substantive revisions of either type do not alter the state of the law as enforced, and are instead designed to aid readability and streamline organization of the Code's various titles and sections. However, because revisions of this kind can involve substantial changes to the

[4]Pub. L. 111-203.

text as written, separating non-substantive from substantive changes to the Code is an important step in my analysis.

To identify non-substantive revisions to the Code, I gathered a list of all major and officially-documented non-substantive Code revision efforts completed from 1995-2016.[5] To develop this list, I drew on three sources. First, the OLRC maintains lists of in-progress and recently-completed editorial reclassification and positive law codification projects. Second, the headnotes to each title in the US Code identifies any and all positive law codification bills which affected the title in question, which I used to gather a list of additional positive-law codification efforts. Third, to check this list for completeness I searched congress.gov's legislation database for any enacted containing the phrases "codify", "United States Code", "positive law", or "without substantive changes" in the title. As shown in Table 5.1, this process yielded some 22 title-year combinations in which a major non-substantive revision to a title of the US code was undertaken in a particular year.

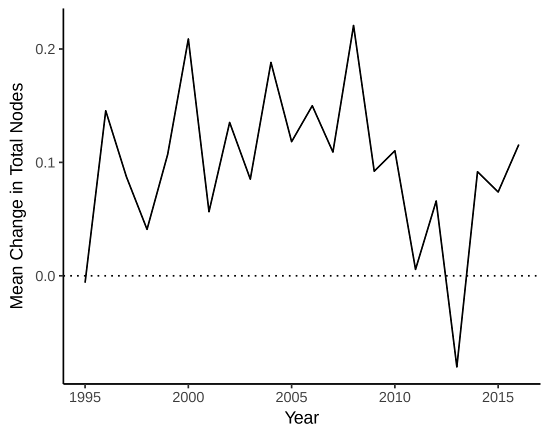
[5]The earliest available online version of Code the is the 1994 edition. However, since all my analyses in this chapter focus on the *differences* between versions of the Code, the 1994 edition of the Code is implicitly excluded from my analysis, allowing me to ignore non-substantive revision efforts that debuted in this version of the Code.

Table 5.1: US Code titles affected by major non-substantive revisions, 1995-2016

Year	Title	Revision Type	Source
1996	49	Positive	Pub. L. 104-287
1997	49	Positive	Pub. L. 105-102
1998	36	Positive	Pub. L. 105-255, Pub. L. 105-354
2002	40	Positive	Pub. L. 107-217
2003	40	Positive	Pub. L. 108-178
2006	46	Positive	Pub. L. 109-304
2010	15	Positive	Pub. L. 111-314[1]
2010	41	Positive	Pub. L. 111-350
2010	42	Positive	Pub. L. 111-314[1]
2010	51	Positive	Pub. L. 111-314[1]
2013	2	Editorial	OLRC
2013	50	Editorial	OLRC
2014	2	Editorial	OLRC[2]
2014	16	Positive	Pub. L. 113-287[3]
2014	36	Positive	Pub. L. 113-237
2014	42	Editorial	OLRC[2]
2014	52	Editorial	OLRC[2]
2014	54	Positive	Pub. L. 113-287[3]
2015	50	Editorial	OLRC
2016	20	Editorial	OLRC[4]
2016	25	Editorial	OLRC
2016	42	Editorial	OLRC[4]

[1] Sections of Titles 15 and 42 were relocated to the newly-created Title 51.
[2] Sections of Titles 2 and 42 were relocated to the newly-created Title 52.
[3] Sections of Title 16 were relocated to the newly-created Title 54.
[4] Sections of Title 42 were relocated to Title 20.

Figure 5.6: Average change in fragmentation over time



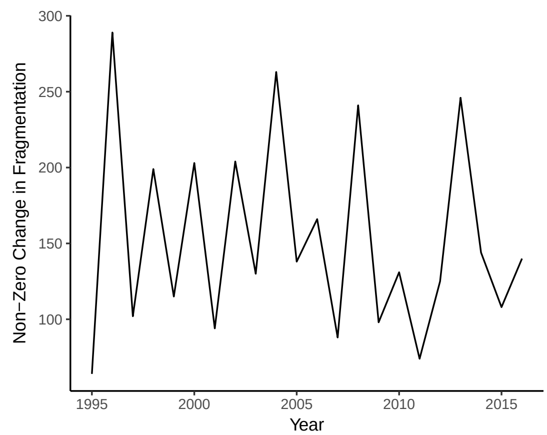
5.3 Variables

5.3.1 The Dependent Variable: Network Fragmentation

Using the chapter-level implementing network dataset described in §5.2.3, I constructed a dataset containing the extracted implementing network for all versions of all US Code chapters from 1994-2016. I then removed all chapters contained in a title affected by a non-substantive revision project (listed in Table 5.1). Since my primary interest in this chapter is in examining the *change* in fragmentation over time, I then removed the first instance in which each chapter appeared in my dataset (usually 1994). Finally, following the procedure I describe in Chapter 4 I operationalize network fragmentation using the *number of unique nodes* present in each network.

In Figures 5.6 and 5.7, I give some basic descriptive information on this dataset.

Figure 5.7: Number of chapters with non-zero change in fragmentation



Overall, the dataset contains some 43,566 chapter-year observations, with 2,534 unique chapters. As shown in Figure 5.6, averaged across all chapters the Code has grown more fragmented in all but two years covered by the dataset. This positive rate of change is unsurprising. Since most new legislation either adds new administrative agencies or allocates new responsibilities to existing ones, we should expect the Code to gradually become more complex over time. However, the average per-chapter change is small, with the largest year-on-year increase (0.224 new unique nodes per chapter) occurring in 2000. Largely, this small effect size reflects the infrequency with which Code chapters are revised; across the dataset, only 8% of chapter-year observations changed in fragmentation from the previous year. As shown in Figure 5.7, this remains roughly stable across the dataset, with approximately 100-300 updated chapters per year.

5.3.2 Predictor Variables

As I describe in §5.1.2, metadata availability for the Code is more limited than for the bills dataset. Only two main chapter-level variables are available: the title from which a particular chapter was drawn, and the date of that chapter. In my modeling work, I use `title` as a rough proxy for policy area and date to generate a dummy variable indicating `unified` government, which mirrors the policy and ideology variables I use in Chapter 4. Because of the difficulty of connecting chapter revisions to their bills, I do not have a way to measure the salience of a particular chapter in a particular year, leaving me without a direct measure of this variable.

Fortunately, however, the availability of multiple Code versions allows me to incorporate a `lagged_fragmentation` variable, which helps alleviate this problem. By controlling for the previous year's fragmentation value, I can control away any time-invariant factors, including the salience of a particular chapter to the extent that it remains constant across the dataset. As I note in §5.3.1, since less than 10% of observations vary from year to year the network fragmentation variable displays approximately a 0.99 year-to-year serial correlation, suggesting that these time-invariant factors explain nearly all of the variance in the dataset. As a result, though the simple lagged term does not capture idiosyncratic year-to-year fluctuations in attention for any given chapter, we can be reasonably confident that it captures most variation in attention in the dataset.

5.4 Modeling

To model network fragmentation in this context, I use a similar Bayesian hierarchical model to the one I employ in Chapter 4. Since my dependent variable is a count, I begin with the following likelihood function:

$$p(y_i|\lambda) = NB(\lambda_i, \phi)$$

Where y_i is the node count value for the i^{th} variable, ϕ represents the scale parameter for the negative binomial distribution and NB represents the negative binomial density. This structure represents a simple negative binomial regression, with λ_i the mean parameter of the regression and ϕ is the scale parameter.

To incorporate covariates in this model, I define:

$$\lambda_i = X_i \beta_{z_i}$$

With X is a $N \times (K + 1)$ matrix of partially-pooled predictor variables and β is a $K \times M$ matrix of partially-pooled coefficients. This structure essentially represents a standard hierarchical regression in which β is allowed to vary by some group structure. In this model, the `unified_government` and `lagged_fragmentation` variables are my only bill-level variables, yielding $K = 2$. Mirroring the model I describe in Chapter 4, I use *bill title* in this chapter as my grouping variable, yielding $M = 52$ groups.

Based on the close correspondence between the `lagged_fragmentation` variable and the current fragmentation variable, I expect that the coefficient on this variable will be positive and significant. I further expect that the coefficient on the `unified` variable will be small for most titles. In Chapter 4, interacting the `unified` variable with a measure of issue salience displayed an attenuation pattern, in which high-salience bills passed under divided/unified government displayed greater differences than their lower-salience counterparts. Theoretically, I would expect a similar pattern to be present in this dataset as well; however, since I cannot directly model salience patterns in the Code, I would expect the overall effect of this variable to be small and frequently non-significant.

As in Chapter 4, I use a hierarchical prior structure to stabilize estimates:

$$\mu_\beta \sim MVN(\beta, \Sigma_\beta)$$

With μ_β a $(K+1)$ vector of top-level regression coefficients and Σ_β a $(K+1)\times(K+1)$ variance-covariance matrix. I place simple independent, mildly informative priors $\mu_k \sim Normal(0, 5)$ on each value in μ, which are intended to constrain the values of these coefficients to plausible ranges.

To complete the model, I place a prior the variance-covariance matrix Σ_β by defining auxiliary variables ν and Ω using the form $\Sigma = diag(\nu)\Omega diag(\nu)$. For numerical stability, I further decompose Ω using a Cholesky factorization such that $\Omega = LL'$, and place the following priors on the auxiliary variables:

$$\nu \sim N_T(0, 10)$$

$$L \sim LKJ(1)$$

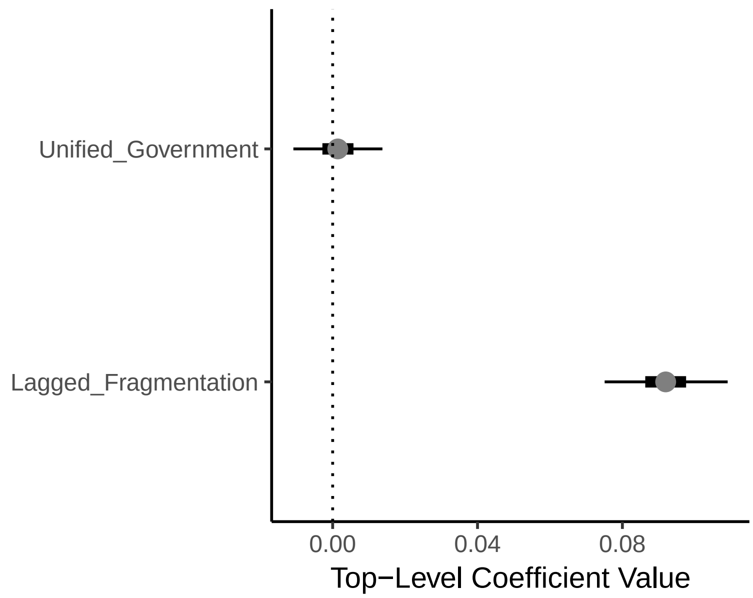
With N_T a half-normal prior, and LKJ denoting the Lewandowski et al. (2009) correlation matrix distribution. $LKJ(1)$ reduces to an identity distribution over correlation matrices, which causes this prior to represent a flat prior over coefficient correlation. As usual, I selected the prior values on ν to represent a vague but mildly informative prior, indicating a slight preference towards coefficient estimates that are smaller in absolute value. In most situations, priors of this kind aid numerical stability during estimation and improve posterior predictive performance.

To fit the model, I used the Stan programming language (Carpenter et al. 2016). I ran four chains, with 750 warmup iterations and 1750 post-warmup iterations in each chain. Visual plots suggested reasonable mixing across chains, with $1 \leq \hat{R} \leq 1.1$ for all parameters and $n_{eff} \geq 500$ for all parameters.[6]

---

[6]With $\hat{R}$ a diagnostic quantifying the consistency of an ensemble of Markov chains, and n_{eff} a rough effective sample size calculation (Gelman et al. 2014). See Appendix B for further details.

Figure 5.8: Top-level non-intercept coefficients

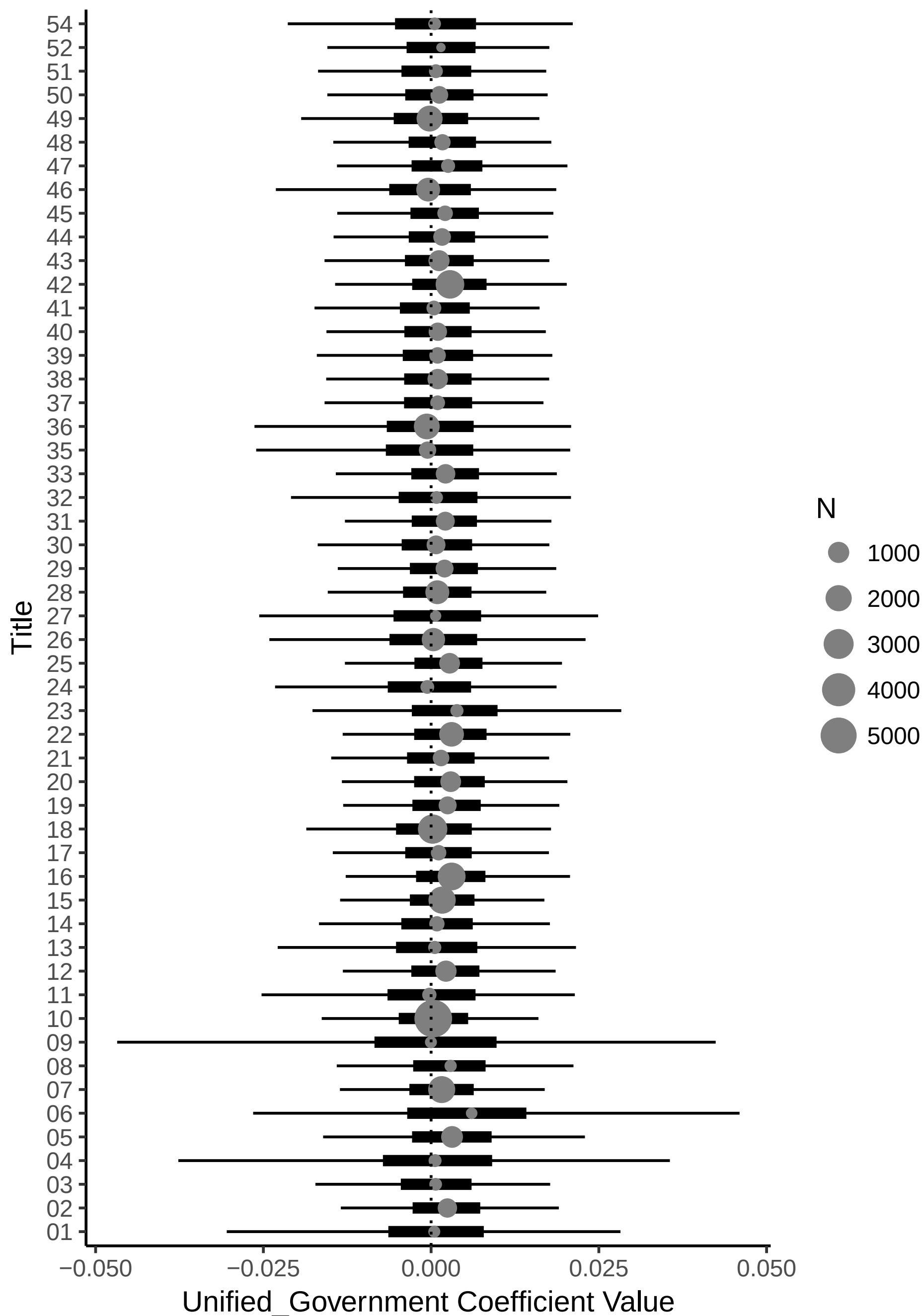


5.5 Results

As with any hierarchical model, a good place to begin in this context is by examining top-level coefficients. As shown in Figure 5.8, the coefficient associated with the `lagged_fragmentation` variable is positive and clearly larger than zero, echoing the clear and strongly positive relationship between the present and lagged dependent variable I noted previously. As in Chapter 4, the coefficient associated with the `unified_government` variable is indistinguishable from zero, suggesting that divided/unified government has a limited impact on fragmentation patterns in the US Code.

As I note in Chapter 4, examining top-level coefficients in a hierarchical model can

Figure 5.9: Lower-level `Unified_Government` coefficients



conceal substantial effect heterogeneity. In Figure 5.9, I show the estimated coefficient values for the `unified_government` variable by title of the US Code. At least in this case, though, the estimated effect of the `unified_government` variable is remarkably consistent across policy areas, with 10% or more of the posterior density both above and below zero for all titles.

As in Chapter 4, this lack of a clear relationship between ideology and downstream design of legislation is surprising, and suggests some interesting directions for future research. One possible explanation for this finding in the Code-based analyses I present in this chapter is my inability to directly control for issue salience, and generating a plausible measure of issue salience for the Code dataset represents an important direction for future work. However, since both the bills dataset and the Code datasets display a minimal relationship between ideology and downstream institutional design choices, it seems likely that the relationship between ideology and design of legislation is at best a weak one.

One other possible explanation for the gap between my findings here and those in studies like Farhang and Yaver (2016) and Epstein and O'Halloran (1999) is the *time-frame* of the datasets under consideration. An array of scholars have demonstrated that post-1994 American politics displays substantially different patterns than the rest of the post-World War 2 era, with sharper ideological divisions and a broader use of "exceptional" procedural tactics. Since datasets like those used by Farhang and Yaver (2016) draw most of their observations from the pre-1994 period, one possible explanation for the divergence between my findings and theirs is this difference in timeframe. Using the methods I develop to examine pre-1994 legislation would be a possible way to probe this difference, and represents an exciting direction for future work.

5.6 Conclusion

Put together, the results I present in this section largely reinforce the findings I describe in Chapter 4. As I describe at the outset of this chapter, a possible problem with the bills-based approach is that it leaves me unable to properly control for *policy context*. Switching from the bills dataset to the US Code leaves me with an impoverished coefficient set; however, in exchange, the versioned nature of the Code allows me to control for policy context by controlling for lagged fragmentation values. Similarly to the bills dataset, the US Code also displays a minimal relationship between divided/unified government and downstream allocation of authority, both on average and within most policy areas.

This analysis leaves a number of directions for future work. In both the bills and the Code datasets, expanding backwards in time would help pinpoint the source of the differences between my results and those reported in related work. In the Code dataset, extracting a parallel salience measure to the one I use in the bills dataset would serve a similar role. One possible direction here is to use the citation notes in each Code chapter to connect the content of that chapter with the corresponding metadata for any cited bills, which would allow me to bridge the gap between these two datasets.

For the time being, however, it seems clear from Chapter 4 and Chapter 5 that ideology is a less significant predictor of downstream institutional design choices than other studies have suggested. Importantly, the results I present should not be read to suggest that executive/legislative preference disagreements have no impact on downstream institutional design choices; instead, my results suggest that, at least for average, day-to-day lawmaking, policy area and issue salience are far more important predictors of administrative structure than ideological differences between the executive and legislative branches. When dealing with high-salience issues and policy areas, ideology may affect legislative behavior more substantially, but at least for the time period I examine the scope of this effect does not appear to be particularly large.

Chapter 6

Conclusion

At the outset of this dissertation, I began with a simple question: how do legislators allocate policymaking authority? This question - or variants thereof - intrigues members of the broader public and has attracted extensive attention in both the legal and political science literatures. Unfortunately, existing public and academic discourse surrounding formal allocation of authority is constrained by serious data availability problems. Both for expert academics and for interested citizens, reading legal texts is difficult, and extracting usable information from these documents is all the more so. As a result, public debate regarding legal texts is often misinformed and the corresponding academic work is limited.

Over the course of this project, I have sought to address these limitations in three parts. First, from a theoretical standpoint, I highlight the gaps that the measurement limitations I identify have caused. As all researchers understand - at least implicitly - measurement structures thought, both by highlighting provocative patterns and by enabling empirical work on measurable questions. In the literature on formal allocation of authority, the labor-intensive nature of existing measurement techniques have generally forced scholars to limit their attention to single policy areas or a small subset of "significant" legislation. However, as I suggest in Chapter 2, based on existing studies of the cognitive dynamics of lawmaker behavior, we should expect these variables to be some of the most important predictors of downstream allocation of authority. By selecting on these variables, existing scholarship therefore misses some of the most important features of the statutory design process.

Second, from a conceptualization and measurement standpoint, I introduce a new scheme designed to capture the institutional features of legal language, which I implement using a machine learning approach. In particular, in Chapter 3 I argue that legal documents are best viewed as *relational* texts, which describe which actors can take which kinds of actions, under which circumstances, and in collaboration with whom. This conceptualization scheme allows researchers to take advantage of mathematical and visualization tools from the network analysis literature, offering a helpful way for researchers to communicate their ideas amongst one another and to members of the public. I then develop, implement, and test a machine learning-based strategy designed to extract implementing networks from legal texts, which I find works well on a pilot dataset of American enacted legislation.

Third, from an empirical standpoint, I use the measurement tools I develop to investigate institutional design patterns in American legislation. In particular, I investigate *fragmentation* of implementing networks in legislative texts. In Chapter 4, I use the tools I develop in this dissertation to extract implementing networks from all American legislation enacted since 1990. I then link this dataset with bill- and legislator-level metadata, and use a hierarchical Bayesian model to examine the relationship between policy area, issue salience, executive-legislative preference disagreements, and fragmentation of each bill's implementing network. As predicted, I find that bills addressing more salient policy areas and policy issues are substantially more fragmented than their counterparts addressing low-salience policy areas and issues. Executive-legislative preference disagreements matter as well, but the effect of this variable is more muted. In some policy areas, high-salience bills passed under unified government are more complex than their unified-government counterparts, but this relationship is relatively weak, limited to a small set of policy areas, and disappears entirely for lower-salience bills.

Finally, in Chapter 5 I supplement this analysis by examining year-on-year changes in fragmentation patterns in the US Code. One potential concern with the bills dataset I examine in Chapter 4 is that it potentially ignores *policy context.* Except in rare cases, bills are usually passed in order to modify some existing administrative or policy structure, rather than to address some novel policy problem. By treating bills as independent units, I therefore ignore the preexisting legal and policy context they address. As a partial solution, I therefore turn to the US Code. Unlike the Statutes at Large, the US Code is versioned, which allows me to isolate variation due to time point-specific factors. After controlling for lagged fragmentation and policy area, I find that fragmentation patterns increase slightly faster under unified than divided government. This effect is consistent across policy areas but substantively small, mirroring my findings in Chapter 4.

The work I present here leaves a number of directions for future research. Within the domain of American legislation, comparing enacted legislation to proposed but unpassed bills would allow me to compare majority- and minority-proposed legislation in a more direct fashion, without selecting for passage. Similarly, comparing the changes in enacted bills across drafts would offer additional insights into the negotiating process underlying each bill, and would allow me to identify the contributions of each individual legislator in a more direct fashion.

Moving beyond the United States, the approach I develop also offer exciting opportunities for comparatively-oriented work. For example, Moe and Caldwell (1994) argue that executive independence from the legislative branch is likely to play a key role in determining the structure of the administrative state. In parliamentary systems, they suggest, because the executive is elected by the legislature the bureaucracy and the legislative branch are likely to posses similar preferences, leading the legislature to pass simple, "framework"-style laws. By contrast, legislators operating in presidential systems are more likely to possess divergent interests from the legislature, and are therefore more likely to pass more restrictive legislation. The measurement techniques I develop in this dissertation offer a way for researchers to test these and related hypotheses directly, without resorting to single policy areas or small sets of "significant" legislation.

Last - but certainly not least - the conceptual and methodological framework I develop in this dissertation provides new avenues for public engagement with legal language. For both experts and non-experts alike, legal language is difficult to read.

As a result, tools that help readers to extract useful language from laws, constitutions, and other legal texts in a quick and efficient fashion have broad applications both within academia and in the broader public discourse. The network-based visualizations and data analysis schemes I present offer one such public engagement tool, which can help inform policy debates beyond the academic domain.

Appendices

Appendix A

Statutory Text Parsing Details

Header Regular Expressions

Table A.1 gives the set of regular expressions used as inputs to the constitute_tools parser, which I use to parse the American legislative text database I introduce in Chapter 3. Note that the list of regular expressions given in this table does not fully capture the set of organizational levels present in American legal language. For example, nearly all major American legislation contains "Title" or variants, which are not captured by this list.

My rationale for excluding some kinds of organizational headers from this list is straightforward. Like many legal corpora, American legislative texts are not written in an entirely consistent fashion. As a result, attempting to capture all organizational headers present across the whole corpus I examine would be highly labor-intensive. Moreover, attempting to extract an organizational header without fully understanding the range of variation present in the corpus can actually do more harm than good; since regular expressions are so flexible, including an overly broad set of regular expressions intended to capture a missing level can potentially consume legally significant language, leading to missing information. By contrast, excluding an organizational header entirely simply introduces a few extraneous words into the set of language under consideration, especially for rarer headers like “Title”.

The selection of regular expressions that I do include in A.1 are included for two reasons. First, as I note in-text, the Office of Law Revision Counsel stipulates that “Sections” of American legislation should be comparable in their substantive scope, which is a structural standard on which I heavily rely throughout my analysis. Segmenting sections is therefore the most critical part of the parsing and text-cleaning task, which is why the regular expression corresponding to “Section” in Table A.1 captures so many special cases and variations. Second, the other headers I include are ubiquitous in nearly every document in the dataset; as a result, excluding them from the list would introduce substantial extraneous language.

Table A.1: Regular expressions used to parse American legislative texts

Regular Expression	Sample Plain-Text Match
(SECTION|SEC\.)\s*\.?\s*(<< NOTE: [0-9]+ USC [-0-9a-z]+\.?\s*’)]	SECTION 101; SEC. 446a
(note)?\.?>>)?\s*[0-9]+\.\s*	344.
\([a-z]\)	(a)
\([0-9]+\)	(34)
\([A-Z]\)	(C)

LSTM Parameter Specification

As described in Chapter 3, I use an LSTM to extract named entities from legislative texts. For the hidden character and word embedding layers, I used a layer sizes of 100 and 300 nodes, respectively. As mentioned in-text, rather than training word embeddings directly I used pre-trained embeddings drawn from Pennington et al. (2014)'s GloVe dataset. Like virtually all neural network applications, I trained this model using stochastic gradient descent.[1] I trained the model for 5 epochs, using 90% of pre-identified named entities for training and 10% as a held-out test set. To avoid overfitting, I use a dropout rate of 0.5 and a batch size of 20, with a learning rate of 0.015, a learning rate decay of 0.05, and a gradient clipping value of 5.0.

[1]Specifically, an ADAM optimizer. See Kingma and Ba (2014) for details.

Appendix B

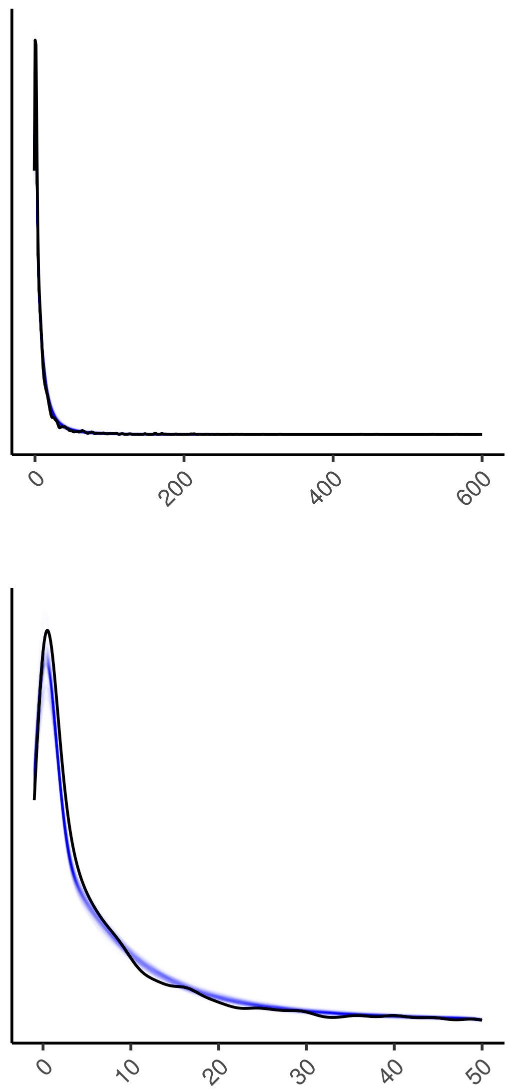
Bayesian Model Details

Enacted Legislation Count Model

Besides the parameter settings given in-text at §4.3, I also initialize all parameter values at 0, and use a maximum treedepth of 15 and an `adapt_delta` value of 0.98. Initial experiments suggested that the default maximum treedepth (10) was sometimes exceeded and a small number of divergent transitions were sometimes encountered with the default adaptation-phase acceptance probability. Increasing these parameters eliminated these problems.

Following Gelman et al. (2014), in Figure B.2 I visually assess model fit using posterior predictive checks. In each plot, I provide the observed density of the node count dependent variable, overlaid on density plots for 400 simulated dependent variable datasets based on randomly-selected post-warmup posterior parameter draws. As shown in the left-hand panel, across the whole dataset the model fit is excellent. Zooming in on smaller values (where most posterior density is located) reveals that the model slightly under-fits at very small values of the dependent variable ($y \leq 10$). Even in this range, however, model fit remains acceptable.

Figure B.1: Posterior predictive plots for the node count dependent variable

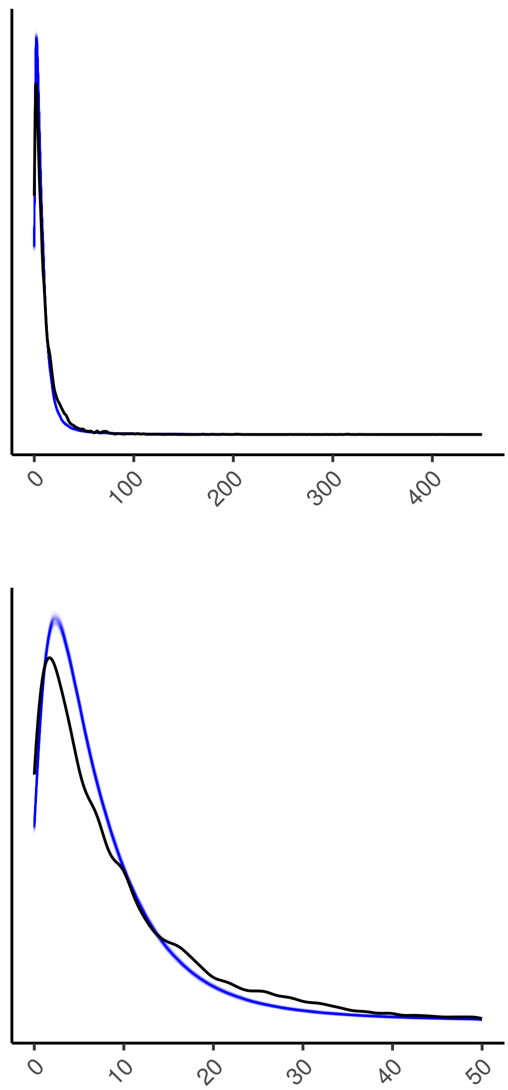


Consolidated Code Count Model

For the Consolidated Code model I describe in §5.4, I initialized all parameter values at zero and used a maximum treedepth of 13. As with the enacted legislation model, I visually assessed model fit using posterior predictive checks. In each plot, I provide the observed density of the node count dependent variable, overlaid on density plots for 400 simulated dependent variable datasets based on randomly-selected post-warmup posterior parameter draws.

Like the bills model, this model underfits the data slightly at low values of the dependent variable, and overfits slightly at large values, but overall, model fit appears acceptable. Rarely, this model also produces unrealistically large predictions on the dependent variable. Across the 7000 post-warmup samples, the largest prediction produced was approximately 1e6, while the largest actually-observed value for the dependent variable was 411. Likely, this divergence is due to the high correlation between the `Lagged_Fragmentation` variable and the dependent variable; since predictions on the dependent variable are so dependent on the estimated coefficient for the `Lagged_Fragmentation` variable, if the model happens to draw an implausibly large value for the coefficient for this variable, the resulting prediction will also be implausibly large. Since model fit overall appears to be good, this pattern is not especially concerning, but it is worth noting for future iterations of this model.

Figure B.2: Posterior predictive plots for the node count dependent variable



Bibliography

Ainsworth, S. H. and Harward, B. M. (2009). Delegation and discretion in anticipation of coalitional drift. *American Politics Research*, 37(6):983–1002.

Anastasopoulos, J. and Bertelli, A. M. (2017). Understanding delegation in the european union through machine learning.

Auerbach, A. J., Gale, W. G., and Harris, B. H. (2010). Activist fiscal policy. *Journal of Economic Perspectives*, 24(4):141–64.

Baumgartner, F. R., Breunig, C., Green-Pedersen, C., Jones, B. D., Mortensen, P. B., Nuytemans, M., and Walgrave, S. (2009). Punctuated equilibrium in comparative perspective. *American Journal of Political Science*, 53(3):603–620.

Baumgartner, F. R. and Jones, B. D. (2010). *Agendas and instability in American politics*. University of Chicago Press.

Bendor, J. and Meirowitz, A. (2004). Spatial models of delegation. *American Political Science Review*, 98(02):293–310.

Bertelli, A. and Feldmann, S. E. (2006). Strategic appointments. *Journal of Public Administration Research and Theory*, 17(1):19–38.

Biber, E. (2011). The more the merrier: Multiple agencies and the future of administrative law scholarship. *Harv. L. Rev. F.*, 125:78.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A. (2016). Stan: A probabilistic programming language. *J Stat Softw.*

Carpenter, D. and Krause, G. A. (2015). Transactional authority and bureaucratic politics. *Journal of Public Administration Research and Theory*, 25(1):5–25.

Clinton, J., Jackman, S., and Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(2):355–370.

de Figueiredo Jr, R. J. (2002). Electoral competition, political uncertainty, and policy insulation. *American Political Science Review*, pages 321–333.

Epstein, D. and O'Halloran, S. (1999). *Delegating powers: a transaction cost politics approach to policy making under separate powers*. Cambridge University Press, New York.

Epstein, S. B. and Kronstadt, K. A. (2013). Pakistan: Us foreign assistance. *Current Politics and Economics of the Middle East*, 4(3):575.

Farhang, S. and Yaver, M. (2016). Divided government and the fragmentation of american law. *American Journal of Political Science*, 60(2):401–417.

Franchino, F. (2004). Delegating powers in the european community. *British Journal of Political Science*, 34(02):269–293.

Franchino, F. (2007). *The powers of the union: delegation in the EU.* Cambridge University Press, New York.

Franzosi, R., De Fazio, G., and Vicari, S. (2012). Ways of measuring agency: an application of quantitative narrative analysis to lynchings in georgia (1875–1930). *Sociological Methodology*, 42(1):1–42.

Gailmard, S. (2002). Expertise, subversion, and bureaucratic discretion. *Journal of Law, Economics, and Organization*, 18(2):536–555.

Gailmard, S. (2009). Multiple principals and oversight of bureaucratic policy-making. *Journal of Theoretical Politics*, 21(2):161–186.

Gailmard, S. and Patty, J. W. (2007). Slackers and zealots: civil service, policy discretion, and bureaucratic expertise. *American Journal of Political Science*, 51(4):873–889.

Gailmard, S. and Patty, J. W. (2012). *Learning while governing: expertise and accountability in the executive branch.* University of Chicago Press, Chicago.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian data analysis.* Taylor & Francis, Boca Raton, FL, 3 edition.

Grimmer, J. and Stewart, B. M. (2013). Text as data: the promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3):267–297.

Huang, Z., Xu, W., and Yu, K. (2015). Bidirectional lstm-crf models for sequence tagging. *arXiv preprint arXiv:1508.01991*.

Huber, J. D. and McCarty, N. (2004). Bureaucratic capacity, delegation, and political reform. *American Political Science Review*, 98(03):481–494.

Huber, J. D. and Shipan, C. R. (2002). *Deliberate discretion? the institutional foundations of bureaucratic autonomy*. Cambridge University Press, New York.

Jones, B. D. and Baumgartner, F. R. (2005). *The politics of attention: How government prioritizes problems*. University of Chicago Press, Chicago, IL.

Jones, B. D. and Baumgartner, F. R. (2012). From there to here: Punctuated equilibrium to the general punctuation thesis to a theory of government information processing. *Policy Studies Journal*, 40(1):1–20.

Kagan, R. A. (2009). *Adversarial legalism: The American way of law*. Harvard University Press.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Koop, C. (2011). Explaining the accountability of independent agencies: The importance of political salience. *Journal of Public Policy*, 31(2):209–234.

Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., and Dyer, C. (2016). Neural architectures for named entity recognition. *arXiv preprint arXiv:1603.01360*.

Leetaru, K. and Schrodt, P. A. (2013). Gdelt: Global data on events, location, and tone, 1979–2012. In *ISA annual convention*, volume 2, pages 1–49. Citeseer.

Lewandowski, D., Kurowicka, D., and Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of multivariate analysis*, 100(9):1989–2001.

Ma, X. and Hovy, E. (2016). End-to-end sequence labeling via bi-directional lstm-cnns-crf. *arXiv preprint arXiv:1603.01354*.

Maltzman, F. and Shipan, C. R. (2008). Change, continuity, and the evolution of the law. *American Journal of Political Science*, 52(2):252–267.

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., and McClosky, D. (2014). The stanford corenlp natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, pages 55–60.

Mayhew, D. R. (1991). *Divided we govern*. Yale University.

McCubbins, M. D. (1985). The legislative design of regulatory structure. *American Journal of Political Science*, pages 721–748.

McCubbins, M. D., Noll, R. G., and Weingast, B. R. (1989). Structure and process, politics and policy: administrative arrangements and the political control of agencies. *Virginia Law Review*, pages 431–482.

McCubbins, M. D. and Schwartz, T. (1984). Congressional oversight overlooked: police patrols versus fire alarms. *American Journal of Political Science*, pages 165–179.

Moe, T. M. (1990a). Political institutions: the neglected side of the story. *Journal of Law, Economics, & Organization*, 6:213–253.

Moe, T. M. (1990b). The politics of structural choice: toward a theory of public bureaucracy. In Williamson, O., editor, *Organization theory: from Chester Barnard to the present and beyond*, pages 116–153. Oxford University Press, New York.

Moe, T. M. (2012). Delegation, control, and the study of public bureaucracy. In *The Forum*, volume 10.

Moe, T. M. and Caldwell, M. (1994). The institutional foundations of democratic government: a comparison of presidential and parliamentary systems. *Journal of Institutional and Theoretical Economics*, 150(1):171–195.

Oberlander, J. (2012). The future of obamacare. *New England Journal of Medicine*, 367(23):2165–2167.

Oosterwaal, A., Payne, D., and Torenvlied, R. (2012). The effect of political disagreement on discretion mechanisms and conditions. *Administration & Society*, 44(7):800–824.

Paas, F., Renkl, A., and Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, 38(1):1–4.

Pennington, J., Socher, R., and Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

Plank, B., Søgaard, A., and Goldberg, Y. (2016). Multilingual part-of-speech tagging with bidirectional long short-term memory models and auxiliary loss. *arXiv preprint arXiv:1604.05529*.

Poole, K. T. and Rosenthal, H. (2000). *Congress: A political-economic history of roll call voting*. Oxford University Press, New York, NY.

Prendergast, C. (2007). The motivation and bias of bureaucrats. *The American Economic Review*, pages 180–196.

Roth, D. and tau Yih, W. (2004). A linear programming formulation for global inference in natural language tasks. In *Proceedings of CoNLL-2004*, pages 1–8. Boston, MA, USA.

Shaffer, R., Whyman, M., Surface Shafran, J., Thomas III, H. F., and Jones, B. D. (2016). Measuring political uncertainty during crises. working paper.

Sinclair, B. (2016). *Unorthodox lawmaking: New legislative processes in the US Congress*. CQ Press.

Stewart, C. I. and Woon, J. (2011). Congressional committee assignments, 103rd to 112th congresses.

Sundermeyer, M., Schlüter, R., and Ney, H. (2012). Lstm neural networks for language modeling. In *Thirteenth Annual Conference of the International Speech Communication Association*.

Surdeanu, M., McClosky, D., Smith, M. R., Gusev, A., and Manning, C. D. (2011). Customizing an information extraction system to a new domain. In *Proceedings of the ACL 2011 Workshop on Relational Models of Semantics*, pages 2–10. Association for Computational Linguistics.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational psychology review*, 22(2):123–138.

Volden, C. (2002). A formal model of the politics of delegation in a separation of powers system. *American Journal of Political Science*, pages 111–133.

Volden, C. and Wiseman, A. E. (2014). *Legislative Effectiveness in the United States Congress: The Lawmakers*. Cambridge University Press.

Vita

Robert Bradley Shaffer was born in Woodland, California, and graduated from the University of California, Berkeley in 2012 with a B.A. in Political Science. He received an M.S. in Statistics and Data Science from the University of Texas, Austin in 2016.

Permanent address: rbshaffer@utexas.edu

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.